# Polynomials and Equations

K.T. Leung I.A.C. Mok S.N. Suen

Polynomials and Equations

# POLYNOMIALS

# AND

# EQUATIONS

K.T. Leung
I.A.C. Mok
S.N. Suen

Hong Kong University Press
香港大學出版社

**Hong Kong University Press**
139 Pokfulam Road, Hong Kong


First published 1992
Reprinted 1993

ISBN 962 209 271 3



Printed in Hong Kong by Nordica Printing Co., Ltd.

# CONTENT


**Preface** vii

**Chapter One Polynomials**

1.1 Terminology 1
1.2 Polynomial functions 4
1.3 The domain $\mathbf{R}[x]$ 8
1.4 Other polynomial domains 14
1.5 The remainder theorem 17
1.6 Interpolation 27

**Chapter Two Factorization of polynomials**

2.1 Divisibility 33
2.2 Divisibility in other polynomial domains 38
2.3 LCM and HCF 41
2.4 Euclidean algorithm 44
2.5 Unique factorization theorem 53

**Chapter Three Notes on the study of equations in ancient civilizations**

3.1 Ancient Egyptian and Babylonian algebra 55
3.2 Ancient Chinese algebra 57
3.3 Ancient Greek algebra 61
3.4 The modern notations 63

**Chapter Four Linear, quadratic and cubic equations**

4.1 Terminology 65
4.2 Linear and quadratic equations 67
4.3 Cubic equations 69
4.4 Equations of higher degree 79

**Chapter Five Roots and coefficients**

5.1 Basic relations 81
5.2 Integral roots 92
5.3 Rational roots 98
5.4 Reciprocal equations 103

**Chapter Six Bounds of real roots**

6.1 The leading term 113
6.2 The constant term 116
6.3 Other bounds of real roots 118

**Chapter Seven The derivative**

7.1 Differentiation 125
7.2 Taylor's formula 132
7.3 Multiple roots 134
7.4 Tangent 140
7.5 Maximum and minimum 142
7.6 Bend points and inflexion points 145

**Chapter Eight Polynomials as continuous functions**

8.1 Continuity 149
8.2 Convergence 153
8.3 Bolzano's theorem 155
8.4 Rolle's theorem 160

**Chapter Nine Separation of real roots**

9.1 The Sturm sequence 167
9.2 Sturm's theorem 171
9.3 Fourier's theorem 180
9.4 Descartes' rule of signs 182

**Chapter Ten Approximation to real roots**

10.1 Newton-Raphson method 187
10.2 Qin-Horner method 192

**Appendix Two theorems on separation of roots** 207

**Numerical answers to exercises** 217

**Index** 231

# PREFACE

Like its predecessor *Fundamental Concepts of Mathematics* (HKUP, 1988) and its successor *Vectors, Matrices and Geometry* (to be published), the present volume *Polynomials and Equations* is primarily a textbook for students of the Sixth Form. It contains the necessary materials for the preparation of the different public examinations of this level in Hong Kong. Moreover, this book also includes parts of the more advanced theory of equations (in Chapters 6, 8, 9 and 10) that are not required in these examinations but are of sufficient importance to serious students of mathematics. Hence it may also serve as a reference book for undergraduate students.

The first two chapters present the algebra of the domain of polynomials with real coefficients and include a proof of the unique factorization theorem which is an important item in the undergraduate algebra syllabus but is usually not required in the Sixth Form examinations. For the benefit of the interested readers, notes are taken, at appropriate places, of polynomials with other coefficients and in more than one indeterminates.

Chapters Three to Five form a self-contained unit on elementary theory of equations. A brief outline of history is given in Chapter Three. This is probably a novelty in a textbook of this level and the section on Chinese mathematics may have a special appeal for students in Hong Kong. Here again Section 4.3 on Cardano's method is some additional material which is not required in the Sixth Form examinations.

In the remaining five chapters of the book polynomials are treated as functions of a real variable. Chapter Seven on derivatives should be relevant to the Sixth Form examination syllabuses. While the derivative of a polynomial is defined here in purely algebraic terms, it is shown to coincide with the analytic notion of the derivative of a differentiable function given in terms of limit. Taylor's expansion is used extensively in the classification of multiple roots. Undergraduate students may find Chapter Eight a useful revision of the most

important concept of continuous function. The results of these two chapters find applications in the separation of roots and in the approximation to roots in the theory of equations presented in the last two chapters of this book.

My former students and friends Miss I.A.C. Mok and Mr. S.N. Suen have provided the book with an excellent set of exercises without which this book would be incomplete and inadequate. My colleagues Dr. M.K. Siu and Dr. K.M. Tsang have been very generous with their suggestions and comments during the preparation of the main text. To them all I would like to express my gratitude. Last but not least I would like to thank Mrs. Annie Cheung for setting the whole text in the present form on AMS-Tex, and Mr. E.T.B. Lau for the line drawings.

K.T. Leung
November 1991

## CHAPTER ONE

# POLYNOMIALS

The study of polynomials constitutes a major component of the mathematics course in secondary school. There polynomials first appear in connection with equations where the main concern is the evaluation of roots. Later they are treated as functions; as such we examine their derivatives, their integrals and their maxima and minima. All along we also learn the arithmetic of polynomials that involves various algebraic operations such as addition, multiplication and factorization of polynomials. In this book we shall continue to study polynomials in these three main aspects.

### 1.1 Terminology

We recall that a *monomial in the indeterminate* $x$ is an expression of the form

$$ax^n$$

where $a$ is a real number and $n$ is a non-negative integer. The real number $a$ is called the *coefficient* and the integer $n$ is called the *exponent* or *power* of $x$ of the monomial $ax^n$. If the coefficient is zero ($a = 0$) then the monomial $ax^n$ is the *zero monomial* and is denoted simply by 0. Therefore all monomials with zero coefficient are identical to the zero monomial: $0x^m = 0x^n = 0$. If $ax^n$ is a non-zero monomial ($a \neq 0$) then the exponent $n$ is called the *degree* of the monomial $ax^n$. By convention the zero monomial 0 shall have no degree. Thus a monomial of degree 0 is a non-zero constant $a$: $ax^0 = a$. It is customary to call monomials of degrees 0, 1, 2 and 3 *constant*, *linear, quadratic* and *cubic* monomials respectively.

Expressions such as $2x^3$, $\frac{1}{\sqrt{5}}\,x$, $2^4$, $\sin\frac{\pi}{7}$, $e$ are monomials whereas expressions such as $|x|$, $\frac{1}{x}$, $\sin x$, $e^x$, $\log x$, $x + x^3$ are not monomials.

Finally two non-zero monomials $ax^n$ and $bx^m$ in the same indeterminate $x$ are equal if and only if they have the same coefficient and the same exponent: $a = b$ and $n = m$. Two non-zero monomials $ax^n$ and $bx^n$ with the same exponent are said to be *alike*. For example, 0 and $\sqrt{2}$ are alike while $x^2$ and $x$ are *unalike*. Two monomials in different indeterminates, e.g. $x$ and $y$, are never equal. Therefore $ax^n \neq by^m$ for whatever coefficients $a$ and $b$, and whatever exponents $n$ and $m$.

A *polynomial in the indeterminate* $x$ is a formal sum of a finite number of unalike monomials. We usually denote polynomials in the indeterminate $x$ by $f(x)$, $g(x)$, $h(x)$, etc. By definition a monomial in $x$ is a polynomial in $x$. A polynomial is usually written as

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

in descending powers of $x$ where the coefficient $a_n$ of the first summand $a_nx^n$ is non-zero. The same polynomial is also written as

$$a_0 + a_1x + \cdots + a_nx^n$$

in ascending powers of $x$. Either way, each summand which is a monomial is called a *term* of the polynomial. The numbers $a_0, a_1, \ldots, a_n$ in the above expressions are called the *coefficients*. The term $a_0$, being a monomial of exponent 0, is called the *constant term* of the polynomial. The term $a_nx^n (a_n \neq 0)$ is called the *leading term*, its coefficient the *leading coefficient* and its degree the *degree* of the polynomial. The degree of $f(x)$ shall be denoted by $\deg f(x)$. Thus every polynomial has a degree which is a non-negative integer except the zero polynomial which, being the zero monomial, has no degree by convention.

Finally two polynomials in $x$

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \qquad (a_n \neq 0)$$

$$g(x) = b_mx^m + b_{m-1}x^{m-1} + \cdots + b_1x + b_0 \qquad (b_m \neq 0)$$

are equal if and only if $n = m$ and $a_i = b_i$ for $i = 0, 1, \ldots, n$. It follows that in writing a polynomial as a sum of monomials it is immaterial in which order its terms appear. Two polynomials in different indeterminates are never equal.

Sometimes it is very important to emphasize the fact that the coefficients $a_i$ of a polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$ are real numbers. To do so, we say that $f(x)$ is a polynomial in $x$ with *real coefficients*, $f(x)$ is a polynomial in $x$ with *coefficients in* $\mathbf{R}$, or $f(x)$ is a polynomial (in $x$) *over* $\mathbf{R}$. The set of all polynomials in $x$ over $\mathbf{R}$ is denoted by $\mathbf{R}[x]$. Here the letter $\mathbf{R}$ indicates that the coefficients are taken from the system $\mathbf{R}$ of real numbers and the letter $x$ indicates the indeterminate under consideration. We shall call $\mathbf{R}[x]$ the *domain of polynomials in $x$ over* $\mathbf{R}$ or the *domain of polynomials in $x$ with coefficients in* $\mathbf{R}$ or the *domain of polynomials with real coefficients*. Obviously $\mathbf{R}$ is a subset of $\mathbf{R}[x]$, since every real number is a monomial of exponent 0.

Monomials and polynomials with coefficients taken from other number systems are similarly defined. Though we shall be mainly concerned with polynomials with real coefficients we may, from time to time, consider polynomials of the domains $\mathbf{Z}[x]$, $\mathbf{Q}[x]$ and $\mathbf{C}[x]$, i.e. polynomials in $x$ whose coefficients are respectively integers, rational numbers and complex numbers.

## EXERCISE 1A

1. Determine whether each of the following polynomials belongs to $\mathbf{Z}[x]$, $\mathbf{Q}[x]$, $\mathbf{R}[x]$ or $\mathbf{C}[x]$:

   (a) $5x^5 - 3x^3 + x$,

   (b) $3x^6 + 4x^2 + \frac{1}{3}x + \sqrt{2}$,

   (c) $\frac{5}{6}x^{11} + \frac{3}{4}x^7 + \frac{1}{2}x^3 - 15$,

   (d) $4x^2 + \sqrt{11}x^{\sqrt{2}} + \frac{1}{5}$,

   (e) $(1 + 3i)x^4 + 5ix - i$, and

   (f) $1 + \sqrt{2}$.

2. A polynomial

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

   is called a monic polynomial if $a_n = 1$.

Let $M[x]$ be the set of monic polynomials in $\mathbf{R}[x]$. Try to construct

(a) a surjective mapping from $M[x]$ to $\mathbf{R}$,

(b) a surjective mapping from $\mathbf{R}[x]$ to $M[x]$, and

(c) an injective mapping from $M[x]$ to $\mathbf{R}[x]$ which does not map any polynomial in $M[x]$ to itself.

## 1.2 Polynomial functions

Given a polynomial

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

of $\mathbf{R}[x]$. If the indeterminate $x$ in the above expression is regarded as a *variable* which can assume any value on the real line $\mathbf{R}$, then in the most natural way the polynomial $f(x)$ will give rise to a mapping of the set $\mathbf{R}$ into $\mathbf{R}$. This mapping is defined by the polynomial $f(x)$ as follows. To each element $c$ of the domain $\mathbf{R}$ there corresponds under the mapping the unique value

$$f(c) = a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0$$

of the range $\mathbf{R}$. Thus it is the mapping $c \to f(c)$ of $\mathbf{R}$ into $\mathbf{R}$. This mapping is called the *polynomial function in the variable $x$ defined by the polynomial* $f(x)$ and shall be denoted also by the same notation $f(x)$ since serious confusion is not likely to occur. For each $c \in \mathbf{R}$ we call the real number $f(c)$ the *value* of the polynomial function (or simply of the polynomial) $f(x)$ at $x = c$.

We remark that though the polynomials

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$
$$g(y) = a_n y^n + a_{n-1}y^{n-1} + \cdots + a_1 y + a_0$$

are unequal as polynomials because they have different indeterminates $x$ and $y$, they define the same polynomial function $c \to f(c) = g(c)$.

The monomial $x$ defines the identity mapping $c \to c$ of $\mathbf{R}$ into $\mathbf{R}$.

If $a$ is constant then the constant polynomial $a$ defines the constant function taking every $c \in \mathbf{R}$ to the constant value $a$. The linear polynomial $v(t) = gt$ in the indeterminate $t$ gives rise to the well-known velocity function of a freely falling body in the variable $t$. Here $t$ measures the time of falling, $g$ is the gravitational constant and $v(t)$ is the velocity of the body at time $t$.

The polynomial functions constitute a very large and important class of functions. First of all they are the simplest type of functions because the value $f(c)$ of a polynomial function $f(x)$ at $x = c$ can be calculated by the elementary algebraic operations of addition and multiplication. Secondly, because of their simplicity, we like to use them to study more complex functions.

Given a polynomial function

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

in the variable $x$, the value $f(c)$ at $x = c$ can be calculated in different ways. For example, we can first calculate successively the powers

$$1, c, c^2, \ldots, c^n$$

of $c$ and then multiply each of them by the corresponding coefficient $a_i$ to get

$$a_0, a_1 c, a_2 c^2, \ldots, a_n c^n$$

and finally we add up to get $f(c)$.

Alternatively we can devise a scheme of *synthetic substitution* based on the following identity with $n-1$ pairs of nested brackets.

$$\begin{aligned} & a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0 \\ = \quad & ((\cdots((a_n c + a_{n-1})c + a_{n-2})c + \cdots + a_2)c + a_1)c + a_0 \ . \end{aligned}$$

According to this identity we may begin with the innermost brackets and calculate step by step as follows.

$$\begin{aligned} & d_n = a_n \\ & d_{n-1} = d_n c + a_{n-1} = a_n c + a_{n-1} \\ & d_{n-2} = d_{n-1} c + a_{n-2} = a_n c^2 + a_{n-1} c + a_{n-2} \\ & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned}$$

$$d_1 = d_2c + a_1 = a_nc^{n-1} + a_{n-1}c^{n-2} + \cdots + a_2c + a_1$$
$$d_0 = d_1c + a_0 = a_nc^n + a_{n-1}c^{n-1} + \cdots + a_2c^2 + a_1c + a_0 = f(c) \ .$$

The synthetic substitution is easily performed on a calculator (of the simplest kind). For calculation by hand the synthetic substitution can be carried out by a *scheme of detached coefficients.* Here we write down three rows of numbers

$$\begin{array}{rcccccccl} & a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 & a_0 & \\ & & d_nc & d_{n-1}c & \cdots & d_3c & d_2c & d_1c & \\ \hline a_n = & d_n & d_{n-1} & d_{n-2} & \cdots & d_2 & d_1 & d_0 & = f(c) \end{array}$$

where the first row contains the coefficients $a_i$ of $f(x)$ and the entries to the third row and the second row are calculated recursively by the recurrent relations

$$d_n = a_n, \qquad d_i = d_{i+1}c + a_i, \qquad d_0 = f(c) \ .$$

The scheme for $f(x) = 5x^3 + 4x^2 - 3$ at $c = -2$ is

$$\begin{array}{lllll} a_3 & a_2 & a_1 & a_0 & \\ & a_3c & a_3c^2 + a_2c & a_3c^3 + a_2c^2 + a_1c & \\ \hline a_3 & a_3c + a_2 & a_3c^2 + a_2c + a_1 & a_3c^3 + a_2c^2 + a_1c + a_0 & = f(c) \end{array}$$

or

$$\begin{array}{rrrrl} 5 & 4 & 0 & -3 & \\ & -10 & 12 & -24 & \\ \hline 5 & -6 & 12 & -27 & = f(-2) \ . \end{array}$$

We notice here that the missing linear term of $f(x)$ is represented by the zero in the top row.

At the special values $c = 0, 1$ and $-1$ the values $f(c)$ is easily seen as

$$\begin{aligned} f(0) &= a_0, \\ f(1) &= a_0 + \cdots + a_n, \\ f(-1) &= a_0 - a_1 + \cdots + (-1)^n a_n \ . \end{aligned}$$

They are respectively the constant term, the sum of the coefficients and the alternating sum of the coefficients of the polynomial $f(x)$.

It is also easy to see that if a non-zero polynomial $f(x)$ has all non-negative or all non-positive coefficients, then $f(c) \neq 0$ for all $c > 0$. On the other hand if the coefficients of $f(x)$ have alternating signs (i.e. either $+, -, +, -, \ldots$ or $-, +, -, +, \ldots$ ), then $f(c) \neq 0$ for all $c < 0$.

## EXERCISE 1B

In what follows, all the polynomials are in $\mathbf{R}[x]$.

1. If $f(x) = 5x^4 + 2x^2 + 3$, evaluate $f(-100)$ and $f(15)$.

2. By using the scheme of detached coefficients, find

   (a) $f(2)$ if $f(x) = x^5 - 6x^4 + 3x^2 + x - 2$, and

   (b) $g(-3)$ if $g(x) = 9x^6 + 10x^5 + 11x^4 + 12x^3 + 13x^2 + 14x + 15$.

   (Part (b) should convince you about the usefulness of the method!)

3. By using the scheme of detached coefficients, find a relation between the real numbers $k$ and $\ell$ where $f(x) = 12x^3 - 17x^2 + kx + \ell$ and $f(-\frac{1}{3}) = 0$. Furthermore, if $f(-\frac{3}{2}) = 0$, find $k$ and $\ell$.

4. The scheme of detached coefficients can also be used to calculate the value $f(c)$ of a polynomial $f(x)$ at some complex number $c$. Now suppose $f(x) = x^4 + 2x^3 - 3x^2 + 4x - 5$. Try to follow the scheme of detached coefficients, find $f(1+i)$ and $f(1-i)$. Verify your answer by substituting $1+i$ and $1-i$ into $f(x)$ directly.

5. Let $f(x) = a_{2n}x^{2n} + a_{2n-2}x^{2n-2} + \cdots + a_2x^2 + a_0$, where $a_i > 0$ for $i = 0, 2, \cdots, 2n-2, 2n$. Show that $f(c) \neq 0$ for every $c \in \mathbf{R}$.

6. Let $f(x) = a_{2n+1}x^{2n+1} + a_{2n-1}x^{2n-1} + \cdots + a_3x^3 + a_1x$, where $a_i > 0$ for $i = 1, 3, \cdots, 2n-1, 2n+1$. Show that $f(c) \neq 0$ for every non-zero $c \in \mathbf{R}$.

7. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a non-zero polynomial. Show that

   (a) if $a_i \geq 0$ or $a_i \leq 0$ for all $i$, then $f(c) \neq 0$ for all real numbers $c > 0$, and

(b) if $a_{2i} \geq 0$ and $a_{2i+1} \leq 0$, or $a_{2i} \leq 0$ and $a_{2i+1} \geq 0$ for $i = 0, 1, \cdots$, then $f(c) \neq 0$ for all real numbers $c < 0$.

## 1.3 The domain $\mathbf{R}[x]$

The domain $\mathbf{R}[x]$ of all polynomials in $x$ with real coefficients contains the set $\mathbf{R}$ of all real numbers as a subset. In $\mathbf{R}$ we have the familiar algebraic operations of addition and multiplication of real numbers; we now want to extend these operations to polynomials so that a useful arithmetic is also available in $\mathbf{R}[x]$.

We define the *sum* $f(x) + g(x)$ of two polynomials

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$
$$g(x) = b_0 + b_1 x + \cdots + b_m x^m$$

of $\mathbf{R}[x]$ to be the polynomial

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

if $m = n$; the polynomial

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m + \cdots + a_n x^n$$

if $m < n$; or the polynomial

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n + \cdots + b_m x^m$$

if $n < m$. Thus the coefficient of $x^i$ in $f(x) + g(x)$ is the sum $a_i + b_i$.

The *product* $f(x)g(x)$ is defined to be the polynomial

$$f(x)g(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \cdots + a_n b_m x^{m+n} .$$

Thus if we denote the product by

$$f(x)g(x) = c_0 + c_1 x + \cdots + c_{n+m} x^{m+n}$$

then the coefficient $c_k$ is given by

$$c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_{k-1} b_1 + a_k a_0 = \sum_{i+j=k} a_i b_j \ .$$

Using a scheme of detached coefficients we can calculate the coefficient $c_k$ as in the following example.

1.3.1 EXAMPLE. *Find the product of* $7x^3+2x^2-4x+7$ *and* $10x^4-5x^2+7$.

SOLUTION:

$$\begin{array}{rrrrrrrr}
 & & & & 7 & 2 & -4 & 7 \\
 & & & 10 & 0 & -5 & 0 & 7 \\
\cline{4-8}
 & & & & 49 & 14 & -28 & 49 \\
 & & -35 & -10 & 20 & -35 & & \\
70 & 20 & -40 & 70 & & & & \\
\hline
70 & 20 & -75 & 60 & 69 & -21 & -28 & 49
\end{array}$$

Thus the product is $70x^7 + 20x^6 - 75x^5 + 60x^4 + 69x^3 - 21x^2 - 28x + 49$.

We take note that for any two polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$, their sum $f(x) + g(x)$ and product $f(x)g(x)$ are both polynomials of the same domain $\mathbf{R}[x]$. We may therefore say that the domain $\mathbf{R}[x]$ is *closed* under the addition and the multiplication defined above. Secondly the sum and the product of two constant polynomials $a$ and $b$ of $\mathbf{R}[x]$ are $a+b$ and $ab$ which are respectively the sum and the product of the real numbers $a$ and $b$ of $\mathbf{R}$. Therefore we may say that the two algebraic operations of $\mathbf{R}[x]$ extend those of $\mathbf{R}$.

It is not difficult to verify that the usual laws of arithmetic hold in $\mathbf{R}[x]$.

The commutative law of addition

$$f(x) + g(x) = g(x) + f(x) \ .$$

The commutative law of multiplication

$$f(x)g(x) = g(x)f(x) \ .$$

The associative law of addition

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) \ .$$

The associative law of multiplication

$$(f(x)g(x))h(x) = f(x)(g(x)h(x)) \ .$$

The distributive law

$$f(x)(g(x) + h(x)) = f(x)g(x) + f(x)h(x) \ .$$

Moreover the constant polynomials 0 and 1 satisfy the following special conditions:

$$f(x) + 0 = f(x) \qquad \text{and} \quad 1f(x) = f(x) \quad \text{for every } f(x) \in \mathbf{R}[x] \ .$$

In fact they are characterized by the above properties.

1.3.2 THEOREM. *$f(x) + g(x) = f(x)$ if and only if $g(x) = 0$. For any non-zero polynomial $f(x)$, $f(x)h(x) = f(x)$ if and only if $h(x) = 1$.*

As in the case of ordinary arithmetic, we may also subtract one polynomial $g(x)$ from another polynomial $f(x)$ to obtain a *difference* $d(x)$ which is itself a polynomial. More precisely, the difference $d(x)$ is given by

$$d(x) = f(x) + (-1)g(x)$$

and is characterized by the property

$$d(x) + g(x) = f(x) \ .$$

Similar to the notation of arithmetic, $(-1)g(x)$ is also written as $-g(x)$ and $d(x) = f(x) + (-1)g(x)$ as $f(x) - g(x)$. Because the difference of two polynomials of $\mathbf{R}[x]$ is again a polynomial of $\mathbf{R}[x]$, we may also say that $\mathbf{R}[x]$ is closed under subtraction.

The division of one polynomial by another is a more complex operation and will be discussed in the next chapter and Section 1.5 of this chapter.

For the degree of the sum and the product we have the useful properties of the theorem below.

1.3.3 THEOREM. *Let $f(x)$ and $g(x)$ be non-zero polynomials of $\mathbf{R}[x]$. Then*

$$\deg(f(x)+g(x)) \leq \max(\deg f(x), \deg g(x))$$
$$\deg f(x)g(x) = \deg f(x) + \deg g(x) \ .$$

PROOF: Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ and $g(x) = b_0 + b_1x + \cdots + b_mx^m$ with $a_n \neq 0$ and $b_m \neq 0$. By definition, $\deg(f(x)+g(x)) = n$ if $m < n$ and $\deg(f(x)+g(x)) = m$ if $n < m$; but $\deg(f(x), g(x)) \leq n$ if $n = m$. Therefore in all cases $\deg(f(x)+g(x)) \leq \max(\deg f(x), \deg g(x))$. It follows from $a_n \neq 0$ and $b_m \neq 0$ that $a_nb_m \neq 0$ in $f(x)g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + a_nb_mx^{n+m}$; hence $\deg f(x)g(x) = \deg f(x) + \deg g(x)$.

We note that in the first general formula of the theorem

$$\deg(f(x)+g(x)) \leq \max(\deg f(x), \deg g(x))$$

the inequality sign cannot be replaced by the equality sign. Take for example $f(x) = 3x^2 + 2x + 1$ and $g(x) = -3x^2 + 5x + 2$. Then $\deg(f(x)+g(x)) = 1 < 2 = \max(\deg f(x), g(x))$. It follows from the second formula of the theorem

$$\deg f(x)g(x) = \deg f(x) + \deg g(x)$$

that if $f(x) \neq 0$ and $g(x) \neq 0$ then $f(x)g(x) \neq 0$. This leads us to formulate the following very simple but very important properties of $\mathbf{R}[x]$.

1.3.4 COROLLARY. *Let $f(x)$ and $g(x)$ be polynomials of $\mathbf{R}[x]$. Then $f(x)g(x) = 0$ if and only if $f(x) = 0$ or $g(x) = 0$.*

1.3.5 COROLLARY. *Let $f(x), g(x)$ and $h(x)$ be polynomials of $\mathbf{R}[x]$. If $f(x) \neq 0$ and $f(x)g(x) = f(x)h(x)$ then $g(x) = h(x)$.*

The last corollary says effectively that we may cancel a non-zero factor from both sides of an equation.

Let us finally consider the algebraic operations in relation to the substitution of values for the indeterminate. Given $f(x)$ and $g(x)$, if

$s(x) = f(x) + g(x)$ and $p(x) = f(x)g(x)$, then for any real number $c$, it is easily verified that

$$s(c) = f(c) + g(c) \qquad \text{and} \qquad p(c) = f(c)g(c) \ .$$

On the left-hand sides of these equations the sum (respectively the product) of the two polynomials is formed prior to the substitution of the value $c$ for $x$. On the right-hand sides the value $c$ is substituted for $x$ before the two values $f(c)$ and $g(c)$ are added (respectively multiplied). In other words we may interchange the order of substitution and algebraic operation without affecting the final result. In terms of polynomial functions, we may also say that the function defined by the sum (product) of two polynomials is identical to the sum (product) of the functions defined by the polynomials.

## EXERCISE 1C

1. Given $f(x) = x^4 + 2x^3 - x^2 - 4x - 2$ and $g(x) = 2x^3 - x^2 + 4$. Find $f(x) + g(x)$, $f(x) - g(x)$ and $f(x)g(x)$.

2. Find real numbers $a$, $b$ and $c$ such that

$$(2x^2 + ax - 1)(x^2 + bx + 1) = 2x^4 + 5x^3 + cx^2 - x - 1 \ .$$

3. Prove that the sum of all the coefficients in the expansion of

$$(8x^9 - 11x^7 + 4x - 2)^3(2x^{10} - 3)^7$$

is 1.

4. Prove that there is no term in odd powers of $x$ of the polynomial

$$f(x) = (x^{100} - x^{99} + x^{98} - x^{97} + \cdots + x^2 - x + 1) \cdot (x^{100} + x^{99} + x^{98} + \cdots + x + 1) \ .$$

(Hint: Try not to expand the product!)

5. Let $f(x)$ be a polynomial of degree $m$, $g(x)$ be a polynomial of degree $n$, $a$ and $b$ are two non-zero real constants. What is the degree of $a \cdot f(x) + b \cdot g(x)$? Give examples to verify your answer.

6. Find $f(x)$ and $g(x)$ of $\mathbf{R}[x]$ such that $\deg f(x) = \deg g(x) = 5$, $\deg(f(x) + g(x)) = 2$ and $\deg(f(x)g(x) + x^5 \cdot f(x)) = 7$.

7. Let $f(x)$ and $g(x)$ be two non-zero polynomials of $\mathbf{R}[x]$.
   (a) What can you say on $g(x)$ if $\deg(f(x)g(x)) = \deg f(x)$?
   (b) What can you say on $f(x)$ and $g(x)$ if $\deg(f(x)g(x)) = 0$?

8. Let $f_1(x)$, $f_2(x)$, $g_1(x)$ and $g_2(x)$ be non-zero polynomials of $\mathbf{R}[x]$. If $\deg f_1(x) \leq \deg g_1(x)$ and $\deg f_2(x) \leq \deg g_2(x)$,
   (a) prove that $\deg f_1(x)f_2(x) \leq \deg g_1(x)g_2(x)$.
   (b) Is it true that

$$\deg(f_1(x) + f_2(x)) \leq \deg(g_1(x) + g_2(x)) \ ?$$

   Justify your answer.

9. Prove the commutative laws of addition and multiplication, the associative laws of addition and multiplication, and the distributive law in $\mathbf{R}[x]$.

10. Prove Theorem 1.3.2.

11. Prove Corollary 1.3.4 and 1.3.5.

12. Given two non-zero polynomials, $h(x)$ and $k(x)$, of $\mathbf{R}[x]$. Let $f(x) = h(x) + (x-a)k(x)$ and $g(x) = (x-a)^m h(x)$, where $a$ is a non-zero real number and $m$ is a positive integer. If $f(x)$ is non-zero and $\deg f(x) < \deg g(x)$, show that $\deg k(x) < \deg h(x) + m - 1$.

13. It is sometimes convenient to give a degree to the zero polynomial 0 by putting $\deg(0) = -\infty$. Show that if the symbol has no other meaning than that $(-\infty) + (-\infty) = -\infty$, $-\infty \leq -\infty$, $-\infty < n$, and $-\infty + n = -\infty$ for all integers $n$, then the two formulae of Theorem 1.3.3 hold also for zero polynomials.

14. Let $c$ be a real number. Define a mapping $\varphi_c : \mathbf{R}[x] \to \mathbf{R}$ by putting $\varphi_c(f(x)) = f(c)$.
   (a) Show that $\varphi_c(f(x) \pm g(x)) = \varphi_c(f(x)) \pm \varphi_c(g(x))$, and $\varphi_c(f(x)g(x)) = \varphi_c(f(x))\varphi_c(g(x))$.
   (b) If $c = 0$, what is the effect of $\varphi_0$ on $f(x)$ of $\mathbf{R}[x]$?
   (c) Is $\varphi_c$ surjective? Justify your answer.

## 1.4 Other polynomial domains

Let us first consider the domain $\mathbf{Z}[x]$ of polynomials in the indeterminate $x$ with coefficients in the number system $\mathbf{Z}$ of integers. A polynomial of the domain $\mathbf{Z}[x]$ is a formal expression of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

where $a_i \in \mathbf{Z}$. Obviously $\mathbf{Z}$ is a subset of $\mathbf{Z}[x]$. On the other hand, since $\mathbf{Z} \subset \mathbf{R}$, we have $\mathbf{Z}[x] \subset \mathbf{R}[x]$, i.e. every polynomial in $x$ over $\mathbf{Z}$ is a polynomial in $x$ over $\mathbf{R}$. Therefore we can speak of *terms, leading coefficients, degree, sum, product*, etc. of polynomials of $\mathbf{Z}[x]$. In particular since $\mathbf{Z}$ is closed under addition and multiplication of integers, the domain $\mathbf{Z}[x]$ is likewise closed under polynomial addition and multiplication. Similarly a polynomial $f(x)$ of $\mathbf{Z}[x]$ defines uniquely a function of $\mathbf{Z}$ into $\mathbf{Z}$ which takes $c \in \mathbf{Z}$ to $f(c) \in \mathbf{Z}$.

Obviously the usual laws of arithmetic hold in $\mathbf{Z}[x]$. The special constant polynomials 1 and 0 also belong to $\mathbf{Z}[x]$ and have the properties given in Theorem 1.3.2 and Corollary 1.3.4.

It would be just a very dull repetition if we were to do the same all over again for $\mathbf{Q}[x]$ and $\mathbf{C}[x]$. It suffices to say that these polynomial domains form a chain of extensions: $\mathbf{Z}[x] \subset \mathbf{Q}[x] \subset \mathbf{R}[x] \subset \mathbf{C}[x]$ and that they all have very similar properties.

Let us consider briefly polynomials in more than one indeterminate. A *monomial* in the indeterminates $x$ and $y$ over $\mathbf{R}$ is an expression of the form

$$ax^m y^n$$

where the coefficient $a$ is a real number and the exponents $m$ and $n$ are non-negative integers. Since $ax^m = ax^m y^0$ and $by^n = bx^0 y^n$, monomials in one indeterminate $x$ or $y$ are monomials in two indeterminates $x$ and $y$. If the monomial $ax^m y^n$ has non-zero coefficient $a$, then the exponent $m$ is the *degree in* $x$, the exponent $n$ the *degree in* $y$ and their sum $m + n$ the *total degree* of the monomial $ax^m y^n$. Two monomials $ax^m y^n$ and $bx^s y^t$ are equal if and only if $a = b$, $m = s$ and $n = t$. Two monomials are *alike* if they have the same exponent in $x$ and the same exponent in $y$.

A *polynomial* in the indeterminates $x$ and $y$ over $\mathbf{R}$ is a formal

sum of a finite number of unalike monomials $a_{rs}x^r y^s$:

$$f(x,y) = \sum a_{rs}x^r y^s \ .$$

For example 3, $x$, $y^2$, $x+y^2$, $5x+xy$ are polynomials in $x$ and $y$. The monomials of the sum are called the *terms* of the polynomial. The various degrees of $f(x,y)$ are the maxima of the corresponding degrees of its non-zero terms. For example, the polynomial $8x^6y^8 - 3x^7y^5 - 5xy^3$ is of degree 7 in $x$ and 8 in $y$, the total degree being 14.

A *homogeneous polynomial* or a *form* is a polynomial in which all terms have the same total degree. For example the polynomial

$$ax + by \qquad \text{where} \quad a \neq 0 \text{ or } b \neq 0$$

is a *linear form* and the non-zero polynomial

$$ax^2 + bxy + cy^2$$

is a *quadratic form.*

Sometimes it is convenient to write down the terms of a polynomial in *lexicographic* order (as in a dictionary) as follows. Of two terms $ax^py^q$ and $bx^ry^s$ with the same total degree (i.e. $p+q=r+s$) $ax^py^q$ preceeds $bx^ry^s$ if $p > r$. Of two terms with different total degrees the one with higher total degree preceeds the other. The *leading term* of a polynomial is the first monomial term (in the lexicographic ordering) among the terms of highest total degree. For example,

$$8x^3 + 5xy^2 + 4y^3 \quad \text{and} \quad 8x^6y^8 + 3x^7y^5 - 2xy^3 + 3x^2 - 3xy$$

are in lexicographic order and their leading terms are $8x^3$ and $8x^6y^8$ respectively.

The set of all polynomials in two indeterminates $x$ and $y$ with real coefficients is called the *domain* of polynomials in $x$ and $y$ over $\mathbf{R}$ and is denoted by $\mathbf{R}[x,y]$. Clearly both $\mathbf{R}[x]$ and $\mathbf{R}[y]$ are subsets of $\mathbf{R}[x,y]$. Addition and multiplication in $\mathbf{R}[x,y]$ are defined similarly; they extend those defined in $\mathbf{R}[x]$ and $\mathbf{R}[y]$ and have similar properties.

Finally a polynomial $f(x,y) = \sum a_{rs}x^r y^s$ defines a function $f(x,y) : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ in two variables $x$ and $y$ which maps an ordered pair $(c,d)$ of real numbers to the real number $f(c,d) = \sum a_{rs}c^r d^s$. It is hardly necessary to repeat for polynomials in two indeterminates over $\mathbf{Z}$, $\mathbf{Q}$ or $\mathbf{C}$.

Polynomial domains in three or more indeterminates are defined similarly. These are denoted by $\mathbf{R}[x,y,z]$, $\mathbf{R}[x_1, x_2, \ldots, x_n]$, etc., where $x_1, x_2, \ldots, x_n$ are $n$ distinct interminates.

## EXERCISE 1D

1. Write down the degree of each of the following polynomials:

   (a) $x^2y^4 + xy^3 - y^4$,

   (b) $10x^7y^3 - 5x^5y^2 + 8x^4y^4 + 4x^6y^6$,

   (c) $3x^2y^3z^4 + 5x^6yz^3 - 7xyz^9$,

   (d) $xyz + x^2y^2z^2 + x^4y^4z^4 - x^2y^4z^6$, and

   (e) $xyzw - x^2yz^2w + 3xy^2z^2w - 5xyz^2w^2$.

2. Rewrite the following polynomial in lexicographic order and pick out the homogeneous polynomials:

   (a) $5x + xy + y^3$,

   (b) $8x^3 + 4y^2 + 5xy^2$,

   (c) $x^2y^2 + xy^3 + x^3y + y^4$,

   (d) $7x^5y^8 + 3x^7y^5 + 2x^4y^9 + 7xy^2 + x^3$, and

   (e) $x^2y^2z + xyz^3 + yz^4 + x^3z^2$.

3. For each of the following pair of polynomials, calculate $f+g$, $f-g$ and $f \cdot g$ and arrange the answers in lexicographic order:

   (a) $f(x,y) = x^2y^2 + 3xy - 2y^2$;
   $g(x,y) = 4x^2y^2 - xy + y^2$;

   (b) $f(x,y) = 5xy + 3x^2y - 4y^2x$,
   $g(x,y) = x^2y + 3xy^2 + 6x^3y$; and

(c) $f(x, y) = 4x^3 + 5y^2 + 8xy^2$,
$g(x, y) = 11x^2y - 6xy^2 + 4y^2 - 5y^3$.

4. Give an example of a polynomial $f(x, y)$ in the indeterminates $x$ and $y$ over $\mathbf{R}$ which satisfies each of the following conditions:

   (a) $f(-x, y) = f(x, y)$,

   (b) $f(x, -y) = f(x, y)$,

   (c) $f(-x, -y) = f(x, y)$, and

   (d) $f(x, y) = f(y, x)$.

5. (a) Show that for each $f(x)$ in $\mathbf{Q}[x]$, there is a positive integer $n_f$ such that $n_f \cdot f(x)$ is in $\mathbf{Z}[x]$, and $n_f$ is the smallest possible one with such property.

   (b) Define a mapping $\varphi : \mathbf{Q}[x] \to \mathbf{Z}[x]$ such that $\varphi(f(x)) = n_f f(x)$ for each $f(x)$ in $\mathbf{Q}[x]$, where $n_f$ is the fixed integer for $f(x)$ as found in (a).

   (i) Show that $\varphi$ is surjective but not injective.

   (ii) Is it true that $\varphi(f(x) + g(x)) = \varphi(f(x)) + \varphi(g(x))$ and $\varphi(f(x) \cdot g(x)) = \varphi(f(x))\varphi(g(x))$ for any $f(x)$, $g(x)$ in $\mathbf{Q}[x]$? Justify your answers.

## 1.5 The remainder theorem

We return to the study of polynomials in one indeterminate. Recall that given a polynomial $f(x)$ of degree $n$ in $\mathbf{R}[x]$, the value $f(c)$ of $f(x)$ at $x = c$ for each $c \in \mathbf{R}$ can be obtained by the method of synthetic substitution according to the following scheme:

$$\begin{array}{ccccccc} & a_n & a_{n-1} & \cdots & a_1 & a_0 & \\ & & d_n c & \cdots & d_2 c & d_1 c & \\ \hline a_n = & d_n & d_{n-1} & \cdots & d_1 & d_0 & = f(c) \end{array}$$

where $a_0, a_1, \ldots, a_n$ are the coefficients of $f(x)$ and the constants $d_0$, $d_1$, $\ldots, d_n$ are formed by the recurrent relations

$$d_n = a_n, \quad d_{i-1} = d_i c + a_{i-1} \ .$$

Explicitly,

$$\begin{aligned}
d_n &= a_n \\
d_{n-1} &= a_n c + a_{n-1} \\
&\cdots\cdots\cdots\cdots\cdots \\
d_1 &= a_n c^{n-1} + a_{n-1}c^{n-2} + \cdots + a_2 c + a_1 \\
d_0 &= a_n c^n + a_{n-1}c^{n-1} + \cdots + a_2 c^2 + a_1 c + a_0 \ .
\end{aligned}$$

Consider the polynomial $q(x)$ of degree $n-1$ defined by

$$q(x) = d_n x^{n-1} + d_{n-1}x^{n-2} + \cdots + d_2 x + d_1$$

where the coefficients are taken from the last row of the scheme of detached coefficients. It follows from the recurrent relations above that

$$\begin{aligned}
(x-c)q(x) &= (x-c)d_n x^{n-1} + (x-c)d_{n-1}x^{n-2} + \cdots + (x-c)d_1 \\
&= d_n x^n + (d_{n-1} - d_n c)x^{n-1} + \cdots + (d_1 - d_2 c)x - d_1 c \\
&= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 - d_0 \\
&= f(x) - f(c) \ .
\end{aligned}$$

Therefore between $f(x)$, $q(x)$, $x-c$ and $f(c)$ the equality

$$f(x) = (x-c)q(x) + f(c)$$

holds. In other words, given any polynomial $f(x)$ of degree $n$ and any number $c$ there exists a polynomial $q(x)$ of degree $n-1$ such that $f(x) = (x-c)q(x) + f(c)$. It is easy to see that the polynomial $q(x)$ is uniquely determined by $f(x)$ and $c$. Suppose $p(x)$ is also a polynomial of degree $n-1$ such that $f(x) = (x-c)p(x) + f(c)$. Then $(x-c)(p(x) - q(x)) = 0$. Since $x - c \neq 0$, by Corollary 1.3.4 we must conclude that $p(x) - q(x) = 0$; hence $p(x) = q(x)$. Thus we have proved the following important theorem.

1.5.1 THE REMAINDER THEOREM. *If $f(x)$ is a polynomial of degree $n \geq 1$ in $\mathbf{R}[x]$ and $c$ is an arbitrary real number, then there exists a unique polynomial $q(x)$ of degree $n-1$ in $\mathbf{R}[x]$ such that*

$$f(x) = (x-c)q(x) + f(c) \ .$$

The polynomial $q(x)$ is called the *quotient* and the constant $f(c)$ the *remainder* of the division of $f(x)$ by $x - c$. This nomenclature is suggested by the following 'long divison' of $f(x)$ by $x - c$:

$$
\begin{array}{llllllll}
 & d_n x^{n-1} & +d_{n-1}x^{n-2} + \cdots\ \cdots & & d_2 x & +d_1 & = & q(x) \\
\hline
(x-c) \mid & a_n x \ +a_{n-1}x^{n-1} & & & +a_1 x & +a_0 & = & f(x) \\
 & a_n x \ - \ d_n c x^{n-1} & & & & & & \\
\cline{2-3}
 & d_{n-1}x^{n-1} & +a_{n-2}x^{n-2} & & & & & \\
 & d_{n-1}x^{n-1} & -d_{n-1}cx^{n-2} & & & & & \\
\cline{2-3}
 & & d_{n-2}x^{n-2} & & & & & \\
 & & & \ddots & & & & \\
 & & & d_2x^2 & +a_1 x & & & \\
 & & & d_2x^2 & -d_2 c x & & & \\
\cline{4-8}
 & & & & d_1 x & +a_0 & & \\
 & & & & d_1 x & -d_1 c & & \\
\cline{5-8}
 & & & & & d_0 & = & f(c)
\end{array}
$$

Therefore $q(x)$ and $f(c)$ do turn out to be the quotient and the remainder of a division. Because of the enormous importance and usefulness of the theorem we felt that it would be instructive to go through another proof to consolidate the idea.

ALTERNATIVE PROOF OF THE REMAINDER THEOREM. We shall not offer another proof for the uniqueness but shall carry out a proof of the existence of the quotient $q(x)$ by induction on the degree $n$ of $f(x)$. For $\deg f(x) = 1$ we have $f(x) = ax + b$ with $a \neq 0$. Putting $q(x) = a$ we get $ax+b = (x-c)a+(ac+b)$ i.e. $f(x) = (x-c)q(x)+f(c)$. Thus the existence of $q(x)$ of degree 0 is proved.

Suppose that for all polynomials of degree less than $n$ such quotients exist. Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \qquad (a_n \neq 0)$$

be a polynomial of degree $n$ and $c$ be a real number. Then the two polynomials $f(x)$ and $(x - c)a_n x^{n-1}$ have identical leading term. Thus

$$g(x) = f(x) - (x - c)a_n x^{n-1}$$

is a polynomial of degree $\leq n-1$. Moreover

$$g(c) = f(c).$$

By induction assumption, there is a quotient $h(x)$ of degree $\leq n-2$ such that

$$g(x) = (x-c)h(x) + g(c).$$

Therefore

$$f(x) - (x-c)a_n x^{n-1} = (x-c)h(x) + f(c).$$

Putting $q(x) = a_n x^{n-1} + h(x)$, which is of degree $n-1$, we have

$$f(x) = (x-c)q(x) + f(c).$$

The induction is now complete.

The remainder theorem provides us with particularly useful information on the polynomial $f(x)$ if the chosen constant $c$ happens to satisfy the condition that $f(c) = 0$, i.e. $c$ is a *zero* (or a *root*) of $f(x)$. In this case we have $f(c) = 0$, and consequently

$$f(x) = (x-c)q(x).$$

Therefore if $c$ is a root of $f(x)$, then the linear polynomial $x-c$ is a *factor* of $f(x)$. Conversely if $x-c$ is a factor of $f(x)$, then $f(x) = (x-c)q(x)$ for some polynomial $q(x)$ whose degree is less than that of $f(x)$ by 1. Recalling that the order of substitution and multiplication can be interchanged, we see that $f(c) = (c-c)q(c) = 0$, i.e. $c$ is a root of $f(x)$. This relationship between a root $c$ and the linear polynomial $x-c$ is known as the factor theorem. Thus we have proved the factor theorem.

1.5.2 THE FACTOR THEOREM. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$ and $c$ be a real number. Then $c$ is a root of $f(x)$ if and only if $x-c$ is a factor of $f(x)$, i.e. $f(x) = (x-c)q(x)$ for some $q(x)$ of $\mathbf{R}[x]$.*

We take note here that the root $c$ of $f(x)$ gets a negative sign in the linear factor $x-c$. Before we proceed to find more useful consequences of the remainder theorem let us work out an example.

1.5.3 EXAMPLE. *Find the roots of the quartic polynomial* $x^4 + x^3 - x - 1$.

SOLUTION: By inspection we find that the sum of the coefficients is zero. Hence $f(1) = 1 + 1 - 1 - 1 = 0$. Therefore 1 is a root of the polynomial $f(x) = x^4 + x^3 - x - 1$. Divide $f(x)$ by $(x - 1)$ to get

$$f(x) = (x - 1)(x^3 + 2x^2 + 2x + 1) .$$

The alternating sum of the coefficients of the polynomial $g(x) = x^3 + 2x^2 + 2x + 1$ is zero. Therefore $-1$ is a root of $g(x)$. Divide $g(x)$ by $(x + 1)$ to get

$$g(x) = (x + 1)(x^2 + x + 1) .$$

Therefore

$$f(x) = (x - 1)(x + 1)(x^2 + x + 1) .$$

The remaining roots of $f(x)$ must be the roots of the quadratic polynomial $h(x) = x^2 + x + 1$. However $h(x)$ has a negative discriminant; it has no real root. Therefore the real roots of $f(x)$ are 1 and $-1$. On the other hand, treating $h(x)$ as a polynomial in $\mathbf{C}[x]$, we see that $h(x)$ has two complex roots $\omega = \frac{1}{2}(-1 + i\sqrt{3})$ and $\omega^2 = \frac{1}{2}(-1 - i\sqrt{3})$ which are primitive cube roots of unity. Therefore $f(x)$ has four complex roots, $-1$ together with the three cube roots of unity.

From this example we see that in $\mathbf{R}[x]$ not only the linear polynomials $x - 1$ and $x + 1$ are factors of the polynomial $f(x)$ but also their product $(x - 1)(x + 1)$. Similarly in $\mathbf{C}[x]$ the linear polynomials $x - 1$, $x + 1$, $x - \omega$, $x - \omega^2$ together with their various products are factors of $f(x)$. This leads us to formulate the following theorem.

1.5.4 THEOREM. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$. If the real (respectively complex) numbers $c_1, c_2, \dots, c_k$ are distinct roots of $f(x)$, thus $f(c_i) = 0$ for $i = 1, 2, \dots, k$, then the $k$-th degree polynomial $(x - c_1)(x - c_2) \cdots (x - c_k)$ is a factor of $f(x)$, i.e.*

$$f(x) = (x - c_1)(x - c_2) \cdots (x - c_k)g(x)$$

*for some polynomial $g(x)$ of $\mathbf{R}[x]$ (respectively of $\mathbf{C}[x]$).*

PROOF: The following inductive proof is based on the factor theorem and the fact that the product of two real (complex) numbers is zero if and only if at least one of the numbers is zero. The induction is carried out on the number $k$ of distinct roots. For $k = 1$ the present theorem is just the factor theorem. Assume that the present theorem holds for $k-1$ distinct roots. Let $c_1, c_2, \ldots, c_k$ be $k$ distinct roots of a polynomial $f(x)$ . Then by induction assumption

$$f(x) = (x - c_2) \cdots (x - c_k)h(x)$$

for some polynomial $h(x)$. Now $f(c_1) = 0$, so $(c_1 - c_2) \cdots (c_1 - c_k)h(c_1) = 0$ since the order of substitution and multiplication may be interchanged. The root $c_1$ is different from $c_2, \ldots, c_k$; therefore $(c_1 - c_2) \cdots (c_1 - c_k) \neq 0$. Hence $h(c_1) = 0$. Applying the factor theorem to $h(x)$ and $c_1$ we have $h(x) = (x - c_1)g(x)$ for some polynomial $g(x)$. Therefore

$$f(x) = (x - c_1)(x - c_2) \cdots (x - c_k)g(x) .$$

We remark that if the roots $c_i$ are not all distinct then the conclusion of the theorem may not hold. Take for example the polynomial $f(x) = x^2 - 1$. $c_1 = -1$ and $c_2 = -1$ are roots of $f(x)$; but $(x+1)^2$ is not a factor of $f(x)$.

One important consequence of this theorem is that a polynomial with $k$ distinct roots must have a degree at least equal to $k$. From this remark a number of corollaries follow.

1.5.5 THEOREM. *A polynomial $f(x)$ of degree $n$ in $\mathbf{R}[x]$ has at most $n$ distinct (real or complex) roots.*

1.5.6 COROLLARY. *If the polynomial expression*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

*with $a_i \in \mathbf{R}$ vanishes upon substitution of $n+1$ distinct values for the indeterminate $x$, then $a_0 = a_1 = \cdots = a_n = 0$.*

1.5.7 COROLLARY. *If two polynomials $f(x)$ and $g(x)$, both of degree $n$ over $\mathbf{R}$, agree in $n+1$ distinct places (i.e. $f(c) = g(c)$ for $n+1$ distinct values of $c$), then they are identical: $f(x) = g(x)$.*

1.5.8 COROLLARY. *If two polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$ are such that $f(c) = g(c)$ for a infinite number of values $c \in \mathbf{R}$ then $f(x) = g(x)$.*

The last corollary may be paraphased as follows. *Distinct polynomials of $\mathbf{R}[x]$ define distinct polynomial functions.*

## EXERCISE 1E

1. If $x+3$ is a factor of $x^3+mx^2+7x+3$, find $m$ and hence factorize the polynomial.
2. Let $f(x) = x - k$ and $g(x) = x^n - k^n$, where $k$ is a real number and $n$ is a positive integer. Find the quotient and remainder when $g(x)$ is divided by $f(x)$.
3. Show that $x-2$ is a factor of $2x^{n+2} - 5x^{n+1} + 2x^n - 2ax^3 + (5a+2)x^2 - (2a+5)x + 2$, where $n$ is a positive integer.
4. By using the same idea as in factor theorem for polynomials of one variable, prove that $x-y$ is a factor of $(y-z)(1+zx)(1+xy) + (z-x)(1+yz)(1+xy) + (x-y)(1+yz)(1+zx)$.
5. Prove that $x-y-z$ is a factor of $x^4+y^4+z^4-2y^2z^2-2z^2x^2-2x^2y^2$.
6. If $3x+1$ and $2x-3$ are factors of $ax^3+bx^2-47x-15$, find the values of $a$ and $b$, and hence factorize the polynomial.
7. Find a polynomial $f(x)$ of degree 2 in $\mathbf{R}[x]$ such that $f(2) = f(3) = 0$, and $f(4) = 6$.
8. Prove Corollary 1.5.6 and Corollary 1.5.7.
9. Find the values of $a$, $b$ and $c$ in each of the following cases so that $f(x) = g(x)$.
   (a) $f(x) = x^2$,
   $g(x) = a(x+1)(x+2) + b(x+1) + c$; and
   (b) $f(x) = (x+1)(x+2)$,
   $g(x) = a(x+3)(x+4) + b(x+5) + c$.
10. Given that $f(x) = x^3 + bx^2 + cx + d$ satisfies the following conditions:
    (a) $f(x)$ has a factor $x-1$,

(b) $f(x)$ has remainder 2 when it is divided by $x-3$, and

(c) $f(x)$ has the same remainder when divided by $x-2$ and $x+2$.

Find $b$, $c$ and $d$.

11. If $a$ and $b$ are two different real numbers, show that $x^2-(a+b)x+ab$ is a factor of

$$x^m(a^n-b^n)+a^m(b^n-x^n)+b^m(x^n-a^n)\ ,$$

where $m$ and $n$ are positive integers.

12. If $(x-a)$ is a common factor of polynomials $f(x)$ and $g(x)$, show that there exist polynomials $p(x)$ and $q(x)$ such that $p(x)f(x)=q(x)g(x)$ with $\deg p(x)<\deg g(x)$ and $\deg q(x)<\deg f(x)$.

13. Find a polynomial $f(x)$ such that $x-1$ is a factor of $f(x)$, and $f(x+h)-f(x)=h(2x+h+1)$ for any real numbers $x$ and $h$.

14. Prove that $a_0-a_2+a_4-\cdots=0$ and $a_1-a_3+a_5-\cdots=0$ are the necessary and sufficient conditions for $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$ to have a factor $x^2+1$.

(Hint: Theorems in this section can be extended from $\mathbf{R}$ to $\mathbf{C}$.)

15. Let $f(x)$ be a polynomial such that

$$f(x)=(x^2-a^2)q(x)+rx+s$$

where $a$, $r$, and $s$ are real numbers with $a\neq 0$, and $q(x)$ is a polynomial.

(a) Show that

$$r=\frac{1}{2a}[f(a)-f(-a)] \qquad \text{and}$$
$$s=\frac{1}{2}[f(a)+f(-a)]\ .$$

In fact, $rx+s$ is the remainder when $f(x)$ is divided by $x^2-a^2$ and $q(x)$ is the quotient. You will learn in the next chapter that when a polynomial $f(x)$ is divided by a non-zero polynomial $g(x)$, the quotient $q(x)$ and remainder $r(x)$ are related by

$$f(x)=g(x)q(x)+r(x)$$

where either $r(x)=0$ or $\deg r(x)<\deg g(x)$ if $r(x)\neq 0$.

Thus, if $\deg g(x) = 2$, $r(x)$ is either a linear polynomial or a constant polynomial.

(b) Hence find the remainder when $x^n - a^n$ is divided by $x^2 - a^2$ when (i) $n$ is even, and (ii) $n$ is odd.

16. Let $n$ be a positive integer greater than 1, and $a \neq -2$ is a real number. If the remainder when $x^n$ is divided by $x^2 + ax - (a+1)$ is $hx + k$, find $h$ and $k$ in terms of $a$.

17. Given that polynomial $f(x)$ gives remainder $3x + 5$ when divided by $(2x+1)(x+2)$. Find the remainders when $f(x)$ is divided by $2x + 1$ and $x + 2$ respectively.

18. Factorize $x^4 + x^3 + x^2 + x$ and hence show that $x^4 + x^3 + x^2 + x$ is a factor of $x^{4444} + x^{3333} + x^{2222} + x^{1111}$.

19. Given $n$ is a positive integer. Find the condition on $n$ such that $x^2+x+1$ is a factor of $x^{2n} + x^n + 1$.

20. Let $q_1(x)$ and $r_1$ be the quotient and remainder when $f(x) = x^8$ is divided by $x+\frac{1}{2}$ respectively. Find the remainder when $q_1(x)$ is divided by $x + \frac{1}{2}$.

21. Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ in $\mathbf{Z}[x]$.

(a) Show that there exists $g(x)$ in $\mathbf{Z}[x]$ such that

$$f(x) - f(k) = (x-k)g(x)$$

for any real number $k$.

(b) Show that $f(10)$ is divisible by 9 if and only if $f(1)$ is divisible by 9.

22. Suppose $f(x)$ is a polynomial and it gives the same remainder when it is divided by $(x-a)(x-b)$ and $(x-a)(x-c)$. Show that $(a-b)f(c) + (b-c)f(a) + (c-a)f(b) = 0$.

23. Show that if $x+r$ is a common factor of $x^3+px^2+q$ and $ax^3+bx+c$, then it is also a factor of $apx^2 - bx + aq - c$. Hence by factorizing a suitable quadratic polynomial, show that $x^3 + \sqrt{7}x^2 - 14\sqrt{7}$ and $2x^3 - 13x - \sqrt{7}$ have a common factor.

24. Given $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, and $\alpha_1, \alpha_2, \ldots, \alpha_n$ be its roots. Find the roots of the following polynomials.

    (a) $a_n x^n - a_{n-1}x^{n-1} + a_{n-2}x^{n-2} - \cdots + (-1)^n a_0$,

    (b) $a_n x^n + a_{n-1}bx^{n-1} + \cdots + a_1 b^{n-1}x + a_0 b^n$, for some real number $b$, and

    (c) $a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$, if all the $\alpha_i \neq 0$.

25. (a) Suppose $p(x)$ is a polynomial with rational coefficients and $\beta = \sqrt[3]{\alpha}$ is an irrational root of $p(x) = 0$, where $\alpha$ is a rational number. Show that $\omega\beta$ and $\omega^2\beta$ are also roots of $p(x) = 0$, where $\omega$ is an imaginary cubic root of unity.

    (b) Find $p(x)$ if $\deg p(x) = 4$ with $p(\sqrt[3]{2}) = 0$, $p(1) = 4$ and $p(0) = 4$.

26. (a) Suppose $f(x)$ is a polynomial with rational coefficients and $a + \sqrt{\beta}$ is an irrational root of $f(x)$, where $\alpha \neq \beta$ are rational numbers. Show that $\alpha - \sqrt{\beta}$ is also a root of $f(x)$.

    (b) Given that $\dfrac{1}{-1-\sqrt{2}}$ is a root of $x^4 - 4x^3 + 5x^2 - 2x - 2$, find the other roots.

27. Let $a_1, a_2, \ldots, a_n$ be $n$ real numbers. If $f(x)$ is a non-zero polynomial of degree $n$ such that

$$f(x) = (x - a_1)q_1(x) + c_0, \qquad c_0 \in \mathbf{R}\ ,$$

$q_{r-1}(x) = (x - a_r)q_r(x) + c_{r-1}$, for $r = 2, 3, \ldots, n$, and $q_n(x) = c_n$, where $q_i(x)$ are polynomials and $c_i$ are real numbers for $i = 1, 2, \ldots, n$.

Find real numbers $d_0, d_1, \ldots, d_n$ such that

$$f(x) = d_n(x-a_1)(x-a_2)\cdots(x-a_n) + d_{n-1}(x-a_1)(x-a_2)\cdots (x-a_{n-1}) + \cdots + d_1(x-a_1) + d_0\ .$$

Now suppose $f(x) = x^5 + x^4 + x^3 + 1$. Rewrite $f(x)$ in the form of

(a) $d_5(x-1)^5 + \cdots + d_1(x-1) + d_0$, and

(b) $d_5(x-1)(x-2)\cdots(x-5) + \cdots + d_1(x-1) + d_0\ .$

## 1.6 Interpolation

We shall consider some more consequences and applications of Corollary 1.5.7.

1.6.1 EXAMPLE. *Show that constants $a, b, c, d$ can be found such that for every integer $n$*

$$n^3 = a(n+1)(n+2)(n+3) + b(n+1)(n+2) + c(n+1) + d\ .$$

SOLUTION: One straightforward way of evaluating $a, b, c, d$ is to (i) expand the right-hand side of the above equation and rewrite it in the standard form of a polynomial in the indeterminate $n$; (ii) equate the corresponding coefficients of the powers of $n$ to get a system of equations in the unknowns $a$, $b$, $c$, $d$; and (iii) solve for $a$, $b$, $c$, $d$. Thus

$$n^3 = an^3 + (6a+b)n^2 + (11a+3b+c) + (6a+2b+c+d)$$

gives rise to the equations

$$a = 1,\ 6a+b = 0,\ 11a+3b+c = 0,\ 6a+2b+c+d = 0\ .$$

We find $a = 1$, $b = -6$, $c = 7$, $d = -1$.

Alternatively we regard

$$\begin{aligned} f(n) &= n^3 \\ g(n) &= a(n+1)(n+2)(n+3) + b(n+1)(n+2) + c(n+1) + d \end{aligned}$$

as cubic polynomials in the indeterminate $n$. In order to apply Corollary 1.5.7 we need 4 distinct values of $n$. Because of the expression $g(n)$, the obvious choice for these values is $-1, -2, -3$ and $0$. Then

$$f(-1) = -1,\ f(-2) = -8,\ f(-3) = -27,\ f(0) = 0\ .$$
$$g(-1) = d,\ g(-2) = -c+d,\ g(-3) = 2b-2c+d,\ g(0) = 6a+2b+c+d\ .$$

Equating each corresponding pair we get

$$d = -1,\ c = 7,\ b = -6,\ a = 1\ .$$

With these values of the constants $a$, $b$, $c$, $d$, the cubic polynomials $f(x)$ and $g(x)$ agree at 4 distinct places. Therefore they are identical.

1.6.2 EXAMPLE. *Prove that for any distinct constants $a, b, c$,*

$$\frac{a^2(x-b)(x-c)}{(a-b)(a-c)} + \frac{b^2(x-c)(x-a)}{(b-c)(b-a)} + \frac{c^2(x-a)(x-b)}{(c-a)(c-b)} = x^2 .$$

PROOF: On each side of the equation above we have a quadratic polynomial. Therefore if they agree at 3 distinct places, they are identical. Clearly they do at $a, b, c$.

The last example suggests an *interpolation method* of finding a polynomial $f(x)$ which should assume a prescribed value $d_i$ at $x = c_i$. Such an interpolation method may have many applications. For example, after an engineer has carried out a series of temperature measurements on a machine at different times, it would be very desirable to be able to express the temperature of the machine as a polynomial function of time.

1.6.3 LAGRANGE'S INTERPOLATION FORMULA. *Given $n+1$ distinct points $c_0, c_1, \dots, c_n$ on the real line $\mathbf{R}$ and $n+1$ arbitrary real values $d_0, d_1, \dots, d_n$, the following formula*

$$f(x) = \sum_{i=0}^{n} \frac{d_i(x-c_0)\cdots(x-c_{i-1})(x-c_{i+1})\cdots(x-c_n)}{(c_i-c_0)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n)}$$

*defines a polynomial $f(x)$ in $\mathbf{R}[x]$ of degree $\leq n$ such that $f(c_i) = d_i$ for $i = 0, 1, ..., n$. Moreover $f(x)$ is the only polynomial of degree $\leq n$ that has this property.*

PROOF: $f(c_i) = d_i$ follows from direct substitution. The uniqueness of $f(x)$ is a consequence of Corollary 1.5.7.

The formula of 1.6.3 is named after Joseph Louis Lagrange (1736–1813). We can also exploit the idea used in the solution of Example 1.6.1 to derive another interpolation formula.

1.6.4 PROBLEM. *Given two sequences of real numbers $c_0, c_1, \dots, c_n, \dots$ and $d_0, d_1, \dots, d_n, \dots$ where the $c_i$ are all distinct from each other. Find a sequence of polynomials $f_0(x), f_1(x), \dots, f_n(x), \dots$ such that for all $n =$ $0, 1, 2, \dots$*

$$\deg f_n(x) \le n \qquad \text{and} \quad f_n(c_i) = d_i \qquad (i \le n)$$

DISCUSSION: The alternative solution to Example 1.6.1 shows that for $c_0 = -1$, $d_0 = -1$; $c_1 = -2$, $d_1 = -8$; $c_2 = -3$, $d_2 = -27$; $c_3 = 0$, $d_3 = 0$; we would obtain

$$\begin{aligned}
f_0 &= -1\\
f_1 &= -1 + 7(x+1)\\
f_2 &= -1 + 7(x+1) - 6(x+1)(x+2)\\
f_3 &= -1 + 7(x+1) - 6(x+1)(x+2) + (x+1)(x+2)(x+3) = x^3
\end{aligned}$$

SOLUTION: We shall find the polynomial $f_n(x)$ one by one beginning with $f_0(x)$. For $n = 0$, we need a constant polynomial because of the requirement on the degree. The obvious choice is $f_0(x) = d_0$. For $n = 1$,

$$f_1(x) = d_0 + \frac{d_1 - d_0}{c_1 - c_0}(x - c_0)$$

will do. Clearly we would need some recursive formulae to define the desired series $f_0(x), f_1(x), \ldots$ of polynomials. Therefore it would be helpful to write

$$\begin{aligned}
f_0(x) =& d_0\\
f_1(x) =& f_0(x) + \frac{d_1 - f_0(c_1)}{c_1 - c_0}(x - c_0) .
\end{aligned}$$

A good guess for the next polynomial would be

$$f_2(x) = f_1(x) + b_2(x - c_0)(x - c_1) .$$

This clearly satisfies the degree requirement that $\deg f_2(x) \le 2$; moreover we have $f_2(c_0) = f_1(c_0) = d_0$ and $f_2(c_1) = f_1(c_1) = d_1$. Threfore what remains to be done is to find the constant $b_2$ such that $f_2(c_2) = d_2$. Writing out $d_2 = f_2(c_2) = f_1(c_2) + b_2(c_2 - c_0)(c_2 - c_1)$, we get

$$b_2 = \frac{d_2 - f_1(c_2)}{(c_2 - c_0)(c_2 - c_1)} .$$

Therefore,

$$f_2(x) = f_1(x) + \frac{d_2 - f_1(c_2)}{(c_2 - c_0)(c_2 - c_1)}(x - c_0)(x - c_1) .$$

It is now not difficult to write down $f_3(x)$ as

$$f_3(x) = f_2(x) + \frac{d_3 - f_2(c_3)}{(c_3 - c_0)(c_3 - c_1)(c_3 - c_2)}(x - c_0)(x - c_1)(x - c_2) .$$

Now $\deg f_3(x) \leq n$, $f_3(c_3) = d_3$ and $f_3(x)$ agrees with $f_2(x)$ at $c_0, c_1$ and $c_2$. The recursive formulae for $f_n(x)$ are given as

$$\begin{aligned} f_0(x) &= d_0 \\ f_n(x) &= f_{n-1}(x) + \frac{d_n - f_{n-1}(c_n)}{(c_n - c_0)\cdots(c_n - c_{n-1})}(x - c_0)\cdots(x - c_{n-1}) . \end{aligned}$$

The formulae are known as Newton's interpolation formulae named after Isaac Newton (1642–1727). While Lagrange's formula has a closed formulation, Newton's formulae are particularly useful if we are dealing with an expanding set of data.

## EXERCISE 1F

1. Find a polynomial $f(x)$ with degree not greater than 3, such that $f(-1) = 3$, $f(0) = 4$, $f(1) = 5$ and $f(2) = 18$.
2. Find a polynomial $f(x)$ of degree not greater than 2 which has the same functional values as $\cos x$ at the points 0, $\frac{\pi}{2}$ and $\pi$.
3. Suppose we were given the information

| $x$ | $-1$ | 0 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | 5 | 2 | 0 | 1 |

   for some unknown function $f(x)$. Try to find $f(1)$ by using Lagrange's interpolation formula.
4. Let $a$, $b$, and $c$ be 3 different real numbers. Find a quadratic polynomial $f(x)$ such that $f(a) = b$, $f(b) = c$ and $f(c) = a$.

   Question 4 can be generalized as follows.
5. Given a sequence of distinct real numbers $a_1, a_2, \ldots, a_{n+1}$, find a polynomial of degree $n$ such that

$$f(a_i) = a_{i+1} \quad \text{for } i = 1, 2, \ldots, n, \text{ and}$$
$$f(a_{n+1}) = a_1 .$$

6. Find the values of $a$, $b$, $c$, and $d$ such that

$$1 \cdot 2 + 2 \cdot 3 + \cdots + n(n+1) = a + bn + cn^2 + dn^3$$

for any positive integer $n$.

7. Prove that for distinct real constants $a_i$, $i = 0, 1, \ldots, n$,

(a) $x^n = \sum_{i=0}^{n} \dfrac{a_i^n(x-a_0)(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_0)(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)}$,

and

(b) $1 = \sum_{i=0}^{n} \dfrac{(x-a_0)(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_0)(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)}$.

8. Theorem 1.6.3 also holds in **C**. Use it to find a linear polynomial $f(x)$ of $\mathbf{C}[x]$ such that $f(i+1) = \frac{1}{2}$ and $f(\pi + i) = \pi$. On the other hand, **Z** does not admit so nice an interpolation theorem. Prove that there is no quadratic polynomial $f(x)$ in $\mathbf{Z}[x]$ satisfying $f(0) = 4$, $f(2) = 6$, and $f(4) = 12$.

   (Hint: Try to consider whether the coefficient of $x$ in $f(x)$ is even or odd.)

   Interpolation property is one of the differences between fields (like **C**, **R**) and domains (like **Z**).

9. To complete the proof of the Newton's interpolation formulae, show that $f_n(c_i) = d_i$ for $i \le n$ and $\deg f_n(x) \le n$ for all $n = 0, 1, 2, \ldots$.

   (a) If $c_k = k$ and $d_k = 2^k$ for $k = 0, 1, 2, \ldots$, find $f_0$, $f_1$, $f_2$ and $f_3$.

   (b) If $c_k = k$ and $d_k = (-1)^k$ for $k = 0, 1, 2, \ldots$, find $f_0$, $f_1$, $f_2$ and $f_3$.

10. Use the Newton's interpolation formulae to do Question 3 again.

CHAPTER TWO

# FACTORIZATION OF POLYNOMIALS

A comparison between the number system $\mathbf{Z}$ and the polynomial domain $\mathbf{R}[x]$ will show that they are very similar as far as formal properties of addition and multiplication are concerned. In fact for most calculations that were carried out in the last chapter, we almost could have operated with polynomials as if they were integers. We shall continue to pursue this similarity in our study of the domain $\mathbf{R}[x]$. In Section 1.5 we have touched upon one very special aspect of factorization of polynomials and found a necessary and sufficient condition for a linear polynomial $(x - c)$ to be a factor of a given polynomial $f(x)$. In the present chapter we shall develop a general theory of factorization in $\mathbf{R}[x]$ aiming at the unique factorization theorem as the counterpart of the fundamental theorem of arithmetic.

## 2.1 Divisibility

In the subsequent discussion we shall tacitly assume that the zero polynomial is excluded and that all polynomials are taken from the domain $\mathbf{R}[x]$. We say that a polynomial $g(x)$ is *divisible* by a polynomial $f(x)$ if $g(x) = f(x)h(x)$ for some polynomial $h(x)$. In this case we also say that $f(x)$ is a *factor* (or a *divisor*) of $g(x)$ or $g(x)$ is a *multiple* of $f(x)$ and write $f(x)|g(x)$.

A non-zero constant polynomial is a factor of every polynomial because for every $a \neq 0$ and every $g(x) = b_m x^m + \cdots + b_1 x + b_0$ we always have $g(x) = a(\frac{b_m}{a}\, x^m + \cdots + \frac{b_1}{a}\, x + \frac{b_0}{a})$. Similarly if $f(x)|g(x)$ then $af(x)|g(x)$ for every non-zero constant $a$. Other general properties of divisibility are listed in the theorem below.

2.1.1 THEOREM. *Let $f(x)$,$g(x)$,$h(x)$,$k(x)$ be polynomials. Then the following statements hold:*

*(a) If $f(x)|g(x)$ and $g(x)|h(x)$, then $f(x)|h(x)$.*
*(b) If $f(x)|g(x)$ and $h(x)|k(x)$, then $f(x)h(x)|g(x)k(x)$.*
*(c) If $f(x)|g(x)$ and $g(x)|f(x)$, then $f(x) = ag(x)$ for some non-zero constant $a$.*
*(d) If $f(x)|g(x)$, then $\deg f(x) \leq \deg g(x)$.*
*(e) If $f(x)|g(x)$ and $f(x)|h(x)$, then $f(x)|(p(x)g(x)+q(x)h(x))$ for arbitrary polynomials $p(x)$ and $q(x)$.*

PROOF: We shall only prove (c) and (d) and leave the proof of the other statements as an exercise.

(c) It follows from the hypothesis that $g(x) = p(x)f(x)$ and $f(x) = q(x)g(x)$ for some polynomials $p(x)$ and $q(x)$. Therefore $f(x) = p(x)q(x)f(x)$; whence $p(x)q(x) = 1$ by Corollary 1.3.5. By Theorem 1.3.3, $\deg p(x) + \deg q(x) = 0$. Therefore both $p(x)$ and $q(x)$ are constant polynomials; so $f(x) = ag(x)$ for some non-zero constant $a$.

(d) If $f(x)|g(x)$ then $f(x)p(x) = g(x)$ for some polynomial $p(x)$ . Therefore $\deg f(x) + \deg p(x) = \deg g(x)$. Since $\deg p(x) \geq 0$, we conclude that $\deg f(x) \leq \deg g(x)$.

Before we study further properties of divisibility, let us compare the statements of the above theorem with their counterparts in the arithmetic of $\mathbf{Z}$ . For non-zero integers $a, b, c$ and $d$ the corresponding statements are:

*(a') If $a|b$ and $b|c$, then $a|c$.*
*(b') If $a|b$ and $c|d$, then $ac|bd$.*
*(c') If $a|b$ and $b|a$, then $a = \pm b$.*
*(d') If $a|b$, then $|a| \leq |b|$.*
*(e') If $a|b$ and $a|c$, then $a|(xb+yc)$ for arbitrary integers $x$ and $y$.*

We discover that (a), (b) and (e) are the exact parallels of (a'), (b') and (e') respectively while there are minor differences between (c) and (c'), and between (d) and (d'). To restore the similarity between (d) and (d'), we can regard the absolute value as a measurement of magnitude of integers and the degree as a measurement of magnitude of polynomials. From this point of view, the statements (d) and (d') are now parallel. Next we observe that 1 and $-1$ are the only integers that have a reciprocal which is also an integer; they are known as the *invertible elements* or *units* of $\mathbf{Z}$. On the other

hand, the invertible polynomials are the non-zero constant polynomials, since they are exactly the polynomials that have a polynomial reciprocal. Therefore we may also call non-zero constant polynomials *units* of $\mathbf{R}[x]$ . The similarity between (c) and (c') is now completely restored:

> *(c) if $f(x)$ and $g(x)$ divide each other, then $f(x) = ag(x)$ for some unit $a$ of $\mathbf{R}[x]$.*
>
> *(c') if $a$ and $b$ divide each other, then $a = cb$ for some unit $c$ of $\mathbf{Z}$.*

If divisibility of integers is our main concern, then we may replace any integer $a$ by $-a$ in a statement about divisibility without altering its validity. Similarly in a statement on divisibility of polynomials, we may replace any polynomial $f(x)$ by any multiple $af(x)$ as long as $a$ is a non-zero constant. This leads us to the following terminology. Two polynomials $f(x)$ and $g(x)$ are said to be *associated* or *associates* of each other if $f(x) = ag(x)$ for some non-zero constant $a$. Among the associates of a given polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ $(a_n \neq 0)$, there is one that has a leading coefficient equal to 1, namely $\frac{1}{a_n} f(x)$. This is called the *monic polynomial associated to* $f(x)$. In general every polynomial with leading coefficient 1 is a *monic* polynomial.

Corresponding to the *prime numbers* of $\mathbf{Z}$ we have the *irreducible polynomials*. Recall that an integer is a *prime number* if it is different from 1 and $-1$, and if it is not a product of two *non-units*. Thus we say that a non-constant polynomial is *irreducible* if it is not a product of two *non-units*, i.e. it is not a product of two polynomials, both of positive degrees. In other words if $f(x)$ is irreducible and $f(x) = g(x)h(x)$ then $\deg g(x) = 0$ or $\deg h(x) = 0$. For example, all linear polynomials are irreducible; so are the quadratic polynmials $x^2 + 1$ and $x^2 + x + 1$, for otherwise they would have a linear factor and hence, by the factor theorem, a root in $\mathbf{R}$, which is impossible. Clearly if $f(x)$ is irreducible and if $f(x)$ and $g(x)$ are associated, then $g(x)$ is also irreducible; in this case the monic polynomial associated to $f(x)$ is irreducible. A non-constant polynomial which is not irreducible is said to be *reducible*; a reducible polynomial can therefore be a product of two non-units, i.e. a product of two polynomials of positive degrees. In other words, if $f(x)$ is reducible, then $f(x) = g(x)h(x)$

for some $g(x)$ and $h(x)$ in $\mathbf{R}[x]$ such that $1 \leq \deg g(x) < \deg f(x)$ and $1 \leq \deg h(x) < \deg f(x)$. For example, $x^2 - 1$, $x^3$, $x^3 - 1$ are reducible.

Like prime numbers, an irreducible polynomial is divisible only by units and its associates.

2.1.2 THEOREM. *Let $p(x)$ be an irreducible polynomial. If $f(x)|p(x)$, then $f(x)$ is either a unit (i.e. a non-zero constant) or an associate of $p(x)$.*

PROOF: It follows from $f(x)|p(x)$ that $p(x) = f(x)g(x)$. Since $p(x)$ is irreducible, either $f(x)$ or $g(x)$ is a unit. In the former case, the theorem holds. In the latter case, $p(x) = af(x)$ for some non-zero constant $a$; hence $p(x)$ and $f(x)$ are associates.

2.1.3 COROLLARY. *Let $p(x)$ and $q(x)$ be irreducible polynomials. If $p(x)|q(x)$, then $p(x)$ and $q(x)$ are associates.*

### EXERCISE 2A

1. Prove (a), (b) and (e) of Theorem 2.1.1.
2. Given polynomials $g(x)$, $f_1(x)$ and $f_2(x)$, show that if
$$g(x)|f_1(x) + f_2(x), \quad g(x)|f_1(x) - f_2(x) \ ,$$
then $g(x)|f_1(x)$ and $g(x)|f_2(x)$.
3. For polynomials $g(x)$, $f_1(x)$ and $f_2(x)$, show that if
$$g(x)|f_1(x), \quad g(x) \nmid f_2(x), \quad \text{then } g(x) \nmid f_1(x) + f_2(x).$$
On the other hand, if $g(x) \nmid f_1(x)$, $g(x) \nmid f_2(x)$, can $f_1(x) + f_2(x)$ be divisible by $g(x)$? Justify your answer.
4. Let $f(x)$ be a polynomial. Show that if $(x-1)|f(x^n)$, then $(x^n - 1)|f(x^n)$.
5. Let $f(x)$ and $g(x)$ be two real polynomials. If $x^2+x+1|f(x^3)+xg(x^3)$, prove that $x-1|f(x)$ and $x-1|g(x)$.

6. Let $a$, $b$, $c$, and $d$ be real numbers such that $abcd \neq 0$. Prove that a necessary and sufficient condition for $ax+b$ to be divisible by $cx+d$ is $\frac{a}{c}=\frac{b}{d}$.

7. If polynomials $f(x)$, $g(x)$ and $h(x)$ satisfy

$$(x^2+1)h(x)+(x-1)f(x)+(x-2)g(x)=0, \quad \text{and}$$
$$(x^2+1)h(x)+(x+1)f(x)+(x+2)g(x)=0\ ,$$

prove that $f(x)$ and $g(x)$ are divisible by $x^2+1$.

8. Prove Corollary 2.1.3.

For questions 9 to 14, $p(x) \sim q(x)$ means $p(x)$ and $q(x)$ are associates.

9. Show that $\sim$ defines an equivalence relation in $\mathbf{R}[x]$. Find the equivalence class of $2x^2+2$.

10. Suppose the non-constant polynomial $p(x)$ is irreducible and $p(x) \sim q(x)$. Show that $q(x)$ is also irreducible.

11. Show that for non-zero polynomials $p(x)$ and $q(x)$, $p(x) \sim q(x)$ if and only if $p(x)|q(x)$ and $q(x)|p(x)$.

12. Let $r(x)$ be a polynomial. For non-zero polynomials $p(x)$ and $q(x)$ such that $p(x) \sim q(x)$, show that
    (a) $r(x)|p(x)$ if and only if $r(x)|q(x)$, and
    (b) $p(x)|r(x)$ if and only if $q(x)|r(x)$.

13. Given non-zero polynomials $p(x)$, $p_1(x)$, $q(x)$ and $q_1(x)$. If $p(x) \sim p_1(x)$ and $q(x) \sim q_1(x)$, show that $p(x)q(x) \sim p_1(x)q_1(x)$. Is it true that $p(x)+q(x) \sim p_1(x)+q_1(x)$? Prove your assertion.

14. Let $a_1$, $a_2$, $\ldots$, $a_n$ be real numbers such that $a_0 \neq 0$ and $a_n \neq 0$. Prove that if $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ is irreducible, then $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ is also irreducible.

## 2.2 Divisibility in other polynomial domains

The definition of divisibility given in the last section can be taken verbatim into the polynomial domains $\mathbf{Q}[x]$ and $\mathbf{C}[x]$. Units in $\mathbf{Q}[x]$ are non-zero rational numbers and units in $\mathbf{C}[x]$ are non-zero complex numbers; they are the non-zero constant polynomials of the respective domains. Therefore statements (a) to (e) of Theorem 2.1.1 hold true in both $\mathbf{Q}[x]$ and $\mathbf{C}[x]$. However more care should be taken with respect to irreducible polynomials since the domains $\mathbf{Q}[x]$, $\mathbf{R}[x]$ and $\mathbf{C}[x]$ have many polynomials in common, and some irreducible polynomials of one domain may become reducible in another domain.

Let us denote by $\mathbf{S}$ a number system which may be $\mathbf{Q}$, $\mathbf{R}$ or $\mathbf{C}$. A non-constant polynomial $f(x)$ of $\mathbf{S}[x]$ is said to be *irreducible over* $\mathbf{S}$ if it is not a product of two non-units of $\mathbf{S}[x]$. By Theorem 1.3.3 all linear polynomials over $\mathbf{S}$ are irreducible over $\mathbf{S}$ irrespectively whether $\mathbf{S}$ is $\mathbf{Q}$, $\mathbf{R}$ or $\mathbf{C}$. For example, $\frac{1}{2}x+1$ is irreducible over $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$.

In general a polynomial which is irreducible over a given number system will remain irreducible over a smaller number system but may fail to be irreducible over a larger number system. Take for example the quadratic polynomials $x^2+1$ and $x^2+x+1$. We have seen that they are irreducible over $\mathbf{R}$ in the last section. They remain irreducible over the smaller number system $\mathbf{Q}$ since they have no root in $\mathbf{Q}$ and hence no linear factor in $\mathbf{Q}[x]$. Both of them become reducible over the larger number system $\mathbf{C}$ as $x^2+1=(x+i)(x-i)$ and $x^2+x+1=(x-\omega)(x-\omega^2)$, where $\omega=\frac{1}{2}(-1+i\sqrt{3})$ is a primitive cube root of unity. Similarly the quadratic polynomials $x^2-2$ and $x^2-2x-4$ are irreducible over $\mathbf{Q}$ but become reducible over $\mathbf{R}$ and hence also over $\mathbf{C}$ as $x^2-2=(x-\sqrt{2})(x+\sqrt{2})$ and $x^2-2x-4=(x-1-\sqrt{5})(x-1+\sqrt{5})$. The observant reader will have noticed that every quadratic polynomial $ax^2+bx+c$ is reducible over $\mathbf{C}$. It is reducible over $\mathbf{R}$ if and only if the discriminant $b^2-4ac$ is non-negative. Finally it is reducible over $\mathbf{Q}$ if and only if the discriminant $b^2-4ac$ is the square of a rational number.

Let us first study the case of $\mathbf{C}[x]$ in detail. The famous fundamental theorem of algebra recapitulated below will provide all information that we need for the study of irreducibility of polynomials of $\mathbf{C}[x]$.

2.2.1 FUNDAMENTAL THEOREM OF ALGEBRA. *A non-constant polynomial of* $\mathbf{C}[x]$ *always has a complex root.*

Consequently every polynomial of $\mathbf{C}[x]$ of degree $\geq 2$ will have a linear factor in $\mathbf{C}[x]$. Therefore *the only irreducible polynomials of* $\mathbf{C}[x]$ *are the linear polynomials.* Moreover, by induction, *every polynomial in* $\mathbf{C}[x]$ *of degree* $n \geq 1$ *is a product of* $n$ *linear factors.* This completes the case study of $\mathbf{C}[x]$.

For a detailed study of polynomials in $\mathbf{R}[x]$ , we need another well-known theorem below.

2.2.2 THEOREM. *Let* $f(x)$ *be a polynomial of* $\mathbf{R}[x]$. *If* $c + di$ *is an imaginary root of* $f(x)$, *then its complex conjugate* $c - di$ *is also an imaginary root of* $f(x)$.

This theorem tells us that the imaginary roots of $f(x) \in \mathbf{R}[x]$ occur in conjugate pairs. On the other hand, if $c + di$ and $c - di$ are a conjugate pair of imaginary numbers ($d \neq 0$), then the product

$$(x - (c + di))(x - (c - di)) = x^2 - 2cx + c^2 + d^2$$

is an irreducible quadratic polynomial of $\mathbf{R}[x]$ with negative discriminant $-4d^2 < 0$. Therefore, if we write $f(x) \in \mathbf{R}[x] \subset \mathbf{C}[x]$ as a product of linear polynomials of $\mathbf{C}[x]$, one for each complex root, then each real root corresponds to a linear factor in $\mathbf{R}[x]$ and each conjugate pair of imaginary roots correspond to an irreducible quadratic factor in $\mathbf{R}[x]$. We can hence conclude that *the only irreducible polynomials of* $\mathbf{R}[x]$ *are the linear polynomials and the quadratic polynomials with negative discriminant.*

The question of irreducibility for polynomials of $\mathbf{Q}[x]$ is far more complex. It suffices here to say that for every $n \geq 1$, there are polynomials in $\mathbf{Q}[x]$ of degree $n$ which are irreducible over $\mathbf{Q}$. For example, $x + 1$, $x^2 + 1$, $x^2 + x + 1$, $x^3 + 4$, $x^4 + 5$ are all irreducible over $\mathbf{Q}$.

Finally let us make a few remarks on the domain $\mathbf{Z}[x]$. The units of the domain $\mathbf{Z}[x]$ are the non-zero constant polynomials 1 and $-1$ since they are the only two invertible elements of $\mathbf{Z}[x]$. Then for

polynomials of $\mathbf{Z}[x]$, statement (c) of Theorem 2.1.1 reads as follows:

*(c) If $f(x)|g(x)$ are $g(x)|f(x)$, then $f(x) = \pm g(x)$*

which is now the same as the statement (c') for $\mathbf{Z}$. Naturally every polynomial irreducible over $\mathbf{Q}$ will remain irreducible over $\mathbf{Z}$. Though $\mathbf{Z}$ is a number system which is smaller than the number system $\mathbf{Q}$, contrary to expectation, every polynomial of $\mathbf{Z}[x]$ which is irreducible over $\mathbf{Z}$ remains irreducible over $\mathbf{Q}$. We shall omit the proof of this classical result as it will carry us too far from the main stream of this course.

## EXERCISE 2B

1. Factorize each of the following polynomials into irreducible factors over (i) $\mathbf{C}$, (ii) $\mathbf{R}$, and (iii) $\mathbf{Q}$.

   (a) $x^3 + 1$

   (b) $x^3 - 3x^2 - 3x + 1$

   (c) $x^4 - x^2 + 2x - 1$

   (d) $x^6 - 1$

2. Show that $x^4 + 1$ is irreducible over $\mathbf{Q}$ but reducible over $\mathbf{R}$.

3. Let $\mathbf{F}$ be $\mathbf{C}$, $\mathbf{R}$ or $\mathbf{Q}$, $a$ is a non-zero element of $\mathbf{F}$ and $f(x) \in \mathbf{F}[x]$.

   (a) If $af(x)$ is irreducible over $\mathbf{F}$, prove that $f(x)$ is irreducible over $\mathbf{F}$.

   (b) If $f(ax)$ is irreducible over $\mathbf{F}$, prove that $f(x)$ is irreducible over $\mathbf{F}$.

   (c) If $f(x+a)$ is irreducible over $\mathbf{F}$, prove that $f(x)$ is irreducible over $\mathbf{F}$.

4. Let $\mathbf{F}$ be $\mathbf{C}$, $\mathbf{R}$ or $\mathbf{Q}$, and $p(x)$ in $\mathbf{F}[x]$ with $\deg p(x) \geq 0$. Show that if $p(x)|f(x) \cdot g(x)$, then $p(x)|f(x)$ or $p(x)|g(x)$ for any $f(x)$, $g(x)$ in $\mathbf{F}[x]$, then $p(x)$ is irreducible over $\mathbf{F}$.

5. If $p(x)$ and $q(x)$ are irreducible polynomials in $\mathbf{R}[x]$ and $p(x)$, $q(x)$ has a common root, show that $p(x)$ and $q(x)$ are associates.

6. We now consider a general criterion for reducibility of quadratic and cubic equations in $\mathbf{F}[x]$, where $\mathbf{F} = \mathbf{C}$, $\mathbf{R}$ or $\mathbf{Q}$.
   If $f(x) \in \mathbf{F}[x]$ and $\deg f(x) = 2$ or $3$, then $f(x)$ is reducible over $\mathbf{F}$ if and only if $f(x)$ has a zero in $\mathbf{F}$.
   Give an example to show that polynomials of degree larger than 3 may be reducible over $\mathbf{F}$ even though they do not have zeros in $\mathbf{F}$.

7. Let $a(x) = a_0 + a_1x + \cdots + a_nx^n$ in $\mathbf{Z}[x]$, and $p$ is a prime number such that

   (i) $p^2 \nmid a_0$

   (ii) $p | a_0, a_1, \ldots, a_{n-1}$

   (iii) $p \nmid a_n$.

   Show that $a(x)$ is irreducible in $\mathbf{Z}[x]$.

8. Using the above result to show that for any prime number $p$ and positive integer $m$, $\sqrt[m]{p}$ is irrational.

## 2.3 LCM and HCF

Let us again return to the study of the domain $\mathbf{R}[x]$ and pursue further its similarity with $\mathbf{Z}$. Given two non-zero integers $a$ and $b$, an integer $d$ is a *greatest common divisor* (gcd for short) of $a$ and $b$ if $d$ is divisor of both $a$ and $b$ and has the greatest absolute value among all common divisors of $a$ and $b$. However $d$ is also characterized in terms of divisibility alone by the two conditions:

(i) $d|a$ and $d|b$,

(ii) if $d'|a$ and $d'|b$ then $d'|d$.

Similarly a *least common multiple* (lcm for short) $m$ is characterized by the two conditions:

(iii) $a|m$ and $b|m$,

(iv) if $a|m'$ and $b|m'$ then $m|m'$.

But $m$ is also a common multiple of $a$ and $b$ with the least absolute value.

Let us carry out the obvious translation.

2.3.1 DEFINITION. Let $f(x)$ and $g(x)$ be non-zero polynomials of $\mathbf{R}[x]$. A polynomial $d(x)$ of $\mathbf{R}[x]$ is a *highest common factor* (HCF for short) of $f(x)$ and $g(x)$ if the following conditions are satisfied:

(i) $d(x)|f(x)$ and $d(x)|g(x)$,

(ii) if $d'(x)|f(x)$ and $d'(x)|g(x)$, then $d'(x)|d(x)$.

A polynomial $m(x)$ of $\mathbf{R}[x]$ is a *lowest common multiple* (LCM for short) of $f(x)$ and $g(x)$ if the following conditions are satisfied:

(iii) $f(x)|m(x)$ and $g(x)|m(x)$,

(iv) if $f(x)|m'(x)$ and $g(x)|m'(x)$ then $m(x)|m'(x)$.

For example, if $f(x) = x^3 - 2x^2 - x + 2 = (x^2 - 1)(x - 2)$ and $g(x) = x^3 + 2x^2 - x - 2 = (x^2 - 1)(x + 2)$, then they have $x^2 - 1$ as an HCF and $x^4 - 5x^2 + 4 = (x^2 - 1)(x^2 - 4)$ as an LCM. The HCF and the LCM of two polynomials are not unique. Clearly any associate of an HCF (respectively an LCM) is an HCF (respectively an LCM) and conversely any two HCFs (respectively LCMs) are associates. We shall usually ignore the distinction between associates and denote any one HCF of $f(x)$ and $g(x)$ by HCF$(f(x), g(x))$ and any one LCM of $f(x)$ and $g(x)$ by LCM$(f(x), g(x))$. We shall prove the existence of HCF and LCM in the next section. In the meantime we proceed to study their properties under the assumption that they exist.

2.3.2 THEOREM. *For any non-zero polynomials $f(x)$, $g(x)$ and $h(x)$, the following statements hold:*

*(a)* $HCF(f(x)h(x), g(x)h(x)) = h(x)HCF(f(x), g(x))$.

*(b)* $LCM(f(x)h(x), g(x)h(x)) = h(x)LCM(f(x), g(x))$.

*(c)* $HCF(f(x), g(x))$ *and* $f(x)$ *are associates if and only if* $f(x)|g(x)$.

*(d)* $LCM(f(x), g(x))$ *and* $f(x)$ *are associates if and only if* $g(x)|f(x)$.

*(e)* $HCF(f(x), g(x)) = HCF(g(x), r(x))$ *if* $f(x) = g(x)q(x) + r(x)$.

*(f)* $HCF(f(x), g(x))LCM(f(x), g(x))$ *is associated to* $f(x)g(x)$.

PROOF: We leave the proof of (a) to (d) to the interested reader as an exercise.

(e) Let $d(x) = HCF(g(x), r(x))$. Then $d(x)|g(x)$ and $d(x)|(g(x)q(x) + r(x))$. Therefore $d(x)|g(x)$ and $d(x)|f(x)$, i.e. condition (i) of Definition 2.3.1 is satisfied. Suppose $d'(x)|f(x)$ and $d'|g(x)$. Then $d'(x)|(f(x) -$

$g(x)q(x))$. Thus $d'(x)|g(x)$, $d'(x)|r(x)$. Since $d(x) = HCF(g(x), r(x))$, we must have $d'(x)|d(x)$, i.e. condition (ii) of Definition 2.3.1, also satisfied. Therefore $d(x) = HCF(f(x), g(x))$. This completes the proof of (e).

(f) Let $m(x) = LCM(f(x), g(x))$. Then it follows from the fact that $f(x)g(x)$ is a common multiple of $f(x)$ and $g(x)$ that $f(x)g(x) = d(x)m(x)$ for some $d(x)$. Therefore it remains to show that $d(x) = HCF(f(x), g(x))$, thus to verify (i) and (ii) of Definition 2.3.1.

Condition (i). It follows from $m(x) = LCM(f(x), g(x))$ that $m(x) = g(x)s(x)$ for some $s(x)$ of $\mathbf{R}[x]$. Then $f(x)g(x) = d(x)m(x) = d(x)g(x)s(x)$. Therefore $f(x) = d(x)s(x)$, and hence $d(x)|f(x)$. Similarly $d(x)|g(x)$.

Condition (ii). Let $d'(x)|f(x)$ and $d'(x)|g(x)$. Then $f(x) = d'(x)h(x)$ and $g(x) = d'(x)k(x)$ for some $h(x)$ and $k(x)$ of $\mathbf{R}[x]$. It follows from $f(x)g(x) = d'(x)h(x)g(x) = d'(x)k(x)f(x)$ that $h(x)g(x) = k(x)f(x)$. Putting $n(x) = h(x)g(x)$, we see that $n(x)$ is a common multiple of $f(x)$ and $g(x)$. Since $m(x) = LCM(f(x), g(x))$, we get $n(x) = m(x)p(x)$. Now it follows from $d(x)m(x) = f(x)g(x) = d'(x)n(x) = d'(x)m(x)p(x)$ that $d(c) = d'(x)p(x)$. Therefore $d'(x)|d(x)$.

Therefore $d(x) = HCF(f(x), g(x))$. This proof of (f) is now complete.

2.3.3 REMARKS. After the obvious modifications are made, Definition 2.3.1 can be used as definitions of HCF and LCM in the other domains $\mathbf{Z}[x]$, $\mathbf{Q}[x]$ and $\mathbf{C}[x]$. Clearly Theorem 2.3.2 also hold in all these domains.

## EXERCISE 2C

1. Prove (a) to (d) of Theorem 2.3.2.

2. For any non-zero polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$, show that

   (a) HCF $(f(x), g(x)) =$ HCF $(f(x) + g(x), g(x))$, and

   (b) HCF $(f(x), g(x)) =$ HCF $(f(x) - g(x), g(x))$.

   (Hint: Apply (e) of Theorem 2.3.2.)

3. By using the results in Question 2, prove that

$$\text{HCF}\left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!},\ 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!}\right) = 1\ .$$

4. For any non-zero polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$, and any polynomial $h(x)$ of $\mathbf{R}[x]$, show that

$$\text{HCF}\,(f(x), g(x)) = \text{HCF}\,(f(x) - h(x)g(x), g(x)) \ .$$

5. For any non-zero polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$, let $f_1(x) = af(x) + bg(x)$, $g_1(x) = cf(x) + dg(x)$, where $a$, $b$, $c$, and $d$ are real numbers such that $ad - bc \neq 0$. Prove that

$$\text{HCF}\,(f(x), g(x)) = \text{HCF}\,(f_1(x), g_1(x)) \ .$$

6. For any non-zero polynomials $f_1(x)$, $f_2(x)$, $g_1(x)$, and $g_2(x)$ of $\mathbf{R}[x]$, prove that

   (a) $\text{HCF}\,(f_1, g_1, f_2, g_2) = \text{HCF}\,(\text{HCF}(f_1, g_1), \text{HCF}(f_2, g_2))$, and

   (b) $\text{HCF}\,(f_1, g_1) \cdot \text{HCF}(f_2, g_2) = \text{HCF}\,(f_1 f_2, f_1 g_2, g_1 f_2, g_1 g_2)$.

   where $f_i$ is an abbreviation of $f_i(x)$ and $g_i$ an abbreviation of $g_i(x)$.

## 2.4 Euclidean algorithm

In the theory of the factorization of integers the following Euclidean algorithm plays a crucial role.

> *Given two non-zero integers $a$ and $b$, there exist unique integers $q$ and $r$ such that*
>
> $$a = bq + r \ \textit{where} \ 0 \leq |r| < |b| \ .$$

Accordingly the division of $a$ by $b$ would either leave no remainder or one whose absolute value is less than that of the divisor $b$. At this stage of our study of polynomials we shall need a similar device in order to make significant progress. Recalling an earlier remark that in the study of divisibility the degree of a polynomial in $\mathbf{R}[x]$ plays the same role as the absolute value of an integer in $\mathbf{Z}$, we have no difficulty in translating the above statement.

2.4.1 EUCLIDEAN ALGORITHM. *If $f(x)$ and $g(x)$ are two non-zero polynomials of $\mathbf{R}[x]$, then there are unique polynomials $q(x)$ and $r(x)$ of $\mathbf{R}[x]$ such that*

$$f(x) = g(x)q(x) + r(x)$$

*where either $r(x) = 0$ or $\deg r(x) < \deg g(x)$ if $r(x) \neq 0$.*

2.4.2 REMARKS. We observe that if $g(x) = x - c$ is a monic linear polynomial, then the above algorithm is just the Remainder Theorem 1.5.1 with $r(x) = f(c)$. In fact the following proof of the present theorem is very similar to the inductive proof of the remainder theorem. The above formulation may appear to be somewhat abstract, it is actually a formalization of the well-known long division of polynomial according to which the division of $f(x)$ by $g(x)$ would either leave no remainder or a remainder $r(x)$ whose degree is less than that of the divisor $g(x)$. Take for example, $f(x) = 6x^5 - 9x^4 + 5x^3 - 20x^2 + 3x - 2$ and $g(x) = 3x^3 - 6x^2 + x - 2$. The long division below

$$\begin{array}{r|llllll}
 & 2x^2+ & x & +3 & & & \\
\hline
3x^3 - 6x^2 + x - 2 & 6x^5 & -9x^4 & +5x^3 & -20x^2 & +3x & -2 \\
 & 6x^5 & -12x^4 & +2x^3 & -4x^2 & & \\
\cline{2-6}
 & & 3x^4 & +3x^3 & -16x^2 & +3x & \\
 & & 3x^4 & -6x^3 & +x^2 & -2x & \\
\cline{3-6}
 & & & 9x^3 & -17x^2 & +5x & -2 \\
 & & & 9x^3 & -18x^2 & +3x & -6 \\
\cline{4-7}
 & & & & x^2 & +2x & +4
\end{array}$$

yields the polynomials $q(x) = 2x^2 + x + 3$ and $r(x) = x^2 + 2x + 4$ which satisfy $f(x) = g(x)q(x) + r(x)$ and $\deg r(x) < \deg g(x)$. It is also because of its close connection with the long division that $q(x)$ and $r(x)$ are called the *quotient* and the *remainder* of the division of $f(x)$ by $g(x)$.

PROOF: The existence. Let

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \qquad (a_n \neq 0)\ ,$$
$$g(x) = b_m x^m + \cdots + b_1 x + b_0 \qquad (b_m \neq 0)\ .$$

If $n < m$ then we can take $q(x) = 0$ and $r(x) = f(x)$. In this case there is nothing more to be proved. Assume that $n \geq m$. We proceed to prove the existence of $q(x)$ and $r(x)$ by induction on $n$. For $n = 0$, we have two non-zero constant polynomials $f(x) = a_0$ and $g(x) = b_0$. In this case we put $q(x) = \frac{a_0}{b_0}$ and $r(x) = 0$. Thus the existence is established. Suppose that for all polynomials $h(x)$ and $g(x)$ such that $n > \deg h(x)$ such quotient and remainder exist. For the given $f(x)$ and $g(x)$ we consider

$$h(x) = f(x) - \frac{a_m}{b_m} x^{n-m} g(x) ,$$

since the two summands on the right-hand side have identical leading term $a_n x^n$, it is clear that $\deg h(x) < n$. By the induction assumption, a quotient $p(x)$ and a remainder $r(x)$ exist such that

$$h(x) = g(x)p(x) + r(x)$$

where either $r(x) = 0$ or $\deg r(x) < \deg g(x)$ if $r(x) \neq 0$. Putting

$$q(x) = \frac{a_n}{b_m} x^{n-m} + p(x) ,$$

we get

$$f(x) = g(x)q(x) + r(x)$$

with $r(x) = 0$ or $\deg r(x) < \deg g(x)$. The induction is complete.

The uniqueness. Suppose we have

$$f(x) = g(x)q(x) + r(x) \quad \text{where} \quad r(x) = 0 \quad \text{or} \quad \deg r(x) < \deg g(x) ,$$

$$f(x) = g(x)q'(x) + r'(x) \quad \text{where} \quad r'(x) = 0 \quad \text{or} \quad \deg r'(x) < \deg g(x) .$$

For the pairs of quotients and remainders we need only consider the case where $q(x) \neq q'(x)$ and $r(x) \neq r'(x)$, the other cases being trivial. It follows from the two equations that $r(x) - r'(x) = g(x)(q'(x) - q(x))$. Since $r(x) - r'(x) \neq 0$ and $q'(x) - q(x) \neq 0$, we have

$$\deg g(x) + \deg(q'(x) - q(x)) \leq \max\{\deg r(x), \deg r'(x)\} < \deg g(x) .$$

But this is impossible since $\deg(q'(x) - q(x)) \geq 0$. Therefore $q(x) = q'(x)$ and $r(x) = r'(x)$.

Having secured the service of the Euclidean algorithm we can now establish the existence of HCF while the existence of LCM would then follow from Theorem 2.3.2(f).

2.4.3. THEOREM. *Two polynomials $f(x)$ and $g(x)$ of $\mathbf{R}[x]$ always have an HCF which can be written in the form*

$$a(x)f(x) + b(x)g(x)$$

*for some polynomials $a(x)$ and $b(x)$ of $\mathbf{R}[x]$.*

PROOF: Consider the set $S = \{s(x)f(x) + t(x)g(x) : s(x), t(x) \in \mathbf{R}[x]\}$. The set $S$ is clearly non-empty and contains non-zero polynomials since both $f(x)$ and $g(x)$ belong to $S$. Among the non-zero polynomials of $S$ we pick any one $d(x)$ which has the lowest degree, say $d(x) = a(x)f(x) + b(x)g(x)$ for certain $a(x)$ and $b(x)$ of $\mathbf{R}[x]$. The theorem will be proved if we can show that $d(x)$ is an HCF of $f(x)$ and $g(x)$. The verification of condition (ii) of 2.3.1 is easy. Since $d(x)$ is of the form $a(x)f(x) + b(x)g(x)$, if $d'(x)|f(x)$ and $d'(x)|g(x)$, then $d'(x)|d(x)$. To show that $d(x)$ satisfies condition (i) of Definition 2.3.1 we shall have to use the device of Euclidean algorithm. Since both $f(x)$ and $g(x)$ belong to $S$, it suffices to show that every polynomial of $S$ is divisible by $d(x)$. Suppose to the contrary there is one element, say $s(x)f(x)+t(x)g(x)$, of $S$ which is not divisible by $d(x)$. Then upon division by $d(x)$, it would leave a non-zero remainder $r(x)$ :

$$s(x)f(x) + t(x)g(x) = d(x)q(x) + r(x)$$

with $\deg r(x) < \deg d(x)$. Then

$$\begin{aligned} r(x) &= s(x)f(x) + t(x)g(x) - d(x)q(x) \\ &= [s(x) - a(x)q(x)]f(x) + [t(x) - b(x)q(x)]g(x) \end{aligned}$$

would be a polynomial of the set $S$ with a degree strictly less than that of $d(x)$. This would contradict our choice of $d(x)$ as a polynomial of $S$ with lowest degree. Therefore $d(x)$ divides every polynomial of $S$ and hence it divides both $f(x)$ and $g(x)$.

Two polynomials $f(x)$ and $g(x)$ are said to be ***relatively prime*** if they have no non-unit common factor, in other words if HCF$(f(x), g(x)) = 1$. Two polynomials being relatively prime is at the one extreme of the possibilities with respective to the availability of common non-unit factors. At the other extreme we would find two polynomials being associates; in this case the polynomials will have all non-unit factors in common. Some of the useful properties of relatively prime polynomials are listed below.

In conjunction with Theorem 2.3.2(f) we have

2.4.4 THEREOM. *$f(x)$ and $g(x)$ are relatively prime if and only if $LCM(f(x), g(x)) = f(x)g(x)$ .*

In conjunction with divisibility we have

2.4.5 THEOREM. *If $f(x)$ and $g(x)$ are relatively prime then $f(x)|h(x)$ follows from $f(x)|g(x)h(x)$.*

PROOF: It follows from Theorem 2.4.3 that $1 = a(x)f(x) + b(x)g(x)$ for some polynomials $a(x)$ and $b(x)$; hence $h(x) = a(x)f(x)h(x) + b(x)g(x)h(x)$. Then both summands on the right-hand side of the last equation are divisible by $f(x)$. Therefore $h(x)$ is divisible by $f(x)$.

In conjunction with irreducible polynomials we have

2.4.6 THEOREM. *If $p(x)$ is an irreducible polynomial, then, for every nonsero polynomial $f(x)$, either $p(x)$ and $f(x)$ are relatively prime or $p(x)|f(x)$.*

PROOF: By hypothesis on $p(x)$ a factor of the irreducible $p(x)$ is either a unit or an associate of $p(x)$. Therefore, either HCF$(p(x), f(x)) = 1$ or HCF$(p(x), f(x)) = p(x)$. In the former case, $p(x)$ and $f(x)$ are relatively prime. In the latter case, $p(x)|f(x)$ by Theorem 2.3.2(c).

2.4.7 COROLLARY. *If $p(x)$ is irreducible and $p(x)|f(x)g(x)$, then $p(x)|f(x)$ or $p(x)|g(x)$.*

PROOF: Consider $p(x)$ and $f(x)$. Then by Theorem 2.4.6 either $p(x)|f(x)$ or $p(x)$ and $f(x)$ are relatively prime. In the former case the corollary holds. In the latter case $p(x)|g(x)$ by Theorem 2.4.5.

By an easy induction we can extend the above corollary to:

2.4.8. COROLLARY. *If $p(x)$ is irreducible and $p(x)|f_1(x)f_2(x)\cdots f_n(x)$, then $p(x)|f_i(x)$ for at least one $f_i(x)$.*

2.4.9. REMARKS. Clearly all the results of this section will hold for the domains $\mathbf{Q}[x]$ and $\mathbf{C}[x]$ without modification. Because division in $\mathbf{Z}$ is not always possible the proof of Euclidean algorithm given above for polynomials of $\mathbf{R}[x]$, which involves division of coefficients, would not be valid for polynomials of $\mathbf{Z}[x]$. In fact, only a weaker form of Euclidean algorithm holds for $\mathbf{Z}[x]$ : *If $f(x)$ and $g(x)$ are non-zero polynomials of $\mathbf{Z}[x]$ and if $b_m$ is the leading coefficient of $g(x)$, then there exist unique polynomials $q(x)$ and $r(x)$ of $\mathbf{Z}[x]$ and a natural number $k$ such that*

$$b_m{}^k f(x) = g(x)q(x) + r(x)$$

*where either $r(x) = 0$ or $\deg r(x) < \deg g(x)$.*

Take for example $f(x) = x^3 + 1$ and $g(x) = 2x + 1$. We find that the algorithm yields

$$2^3(x^3 + 1) = (2x + 1)(4x^2 - 2x + 1) + 7$$

with $q(x) = 4x^2 - 2x + 1, r(x) = 7$ and $k = 3$. The interested reader may like to carry out the necessary modification to the proof of Euclidean algorithm 2.4.1 as an exercise.

## EXERCISE 2D

In what follows, all the polynomials are in $\mathbf{F}[x]$, where $\mathbf{F} = \mathbf{C}$, $\mathbf{R}$ or $\mathbf{Q}$.

1. The Euclidean algorithm can be used to find the HCF of any two non-zero polynomials $f(x)$, $g(x)$.

By carring out the division process a finite number of times, we have

$$
\begin{array}{ll}
f(x) = g(x)q_1 + r_1(x) & \deg r_1(x) < \deg g(x) \text{ or } r_1(x) = 0 \\
g(x) = r_1(x)q_2(x) + r_2(x) & \deg r_2(x) < \deg r_1(x) \text{ or } r_2(x) = 0 \\
\vdots & \\
r_i(x) = r_{i+1}(x)q_{i+2}(x) + r_{i+2}(x) & \deg r_{i+2}(x) < \deg r_{i+1}(x) \\
 & \text{ or } r_{i+2}(x) = 0 \\
\vdots & \\
r_{n-1}(x) = r_n(x)q_{n+1}(x) . &
\end{array}
$$

Show that $r_n(x) = \text{HCF}\,(f(x), g(x))$.

2. Find HCF of each of the following pairs of polynomials.

   (a) $f(x) = 3x^4 + 8x^2 - 3$, and
   $g(x) = x^3 + 2x^2 + 3x + 6$.

   (b) $f(x) = x^6 - x^5 + x^4 - 2x^3 + x^2 - x + 1$, and
   $g(x) = x^5 - 2x^4 + x^3 - x^2 + 2x - 1$.

3. Find the HCF and LCM of $2x^4 + 9x^3 + 14x + 3$ and $3x^4 + 15x^3 + 5x^2 + 10x + 2$.

4. For each of the following pairs of polynomials $f(x)$, $g(x)$, find polynomials $a(x)$, $b(x)$ of $\mathbf{R}[x]$ with the least possible degrees such that

$$a(x)f(x) + b(x)g(x) = 1 .$$

   (a) $f(x) = x^3 - 2x^2 + x - 1$, and
   $g(x) = x^2 + x - 3$.

   (b) $f(x) = x^3 - 3x + 1$, and
   $g(x) = x^2 + x + 1$.

   What can you say about each pair of $f(x)$ and $g(x)$?

5. If $f(x)$ of $\mathbf{F}[x]$ is irreducible over $\mathbf{F}$ and for any polynomial $g(x)$ of $\deg g(x) < \deg f(x)$, show that $f(x)$ and $g(x)$ are relatively prime.

6. If $\text{HCF}\,(f(x), g(x)) = d(x)$ and $f(x) = d(x)m(x)$, $g(x) = d(x)n(x)$, show that $m(x)$ and $n(x)$ are relatively prime.

7. If non-zero polynomials $f(x)$ and $g(x)$ are relatively prime and $r(x)f(x) = s(x)g(x)$, for some polynomials $r(x)$ and $s(x)$, show that $f(x)|s(x)$ and $g(x)|r(x)$.

8. If non-zero polynomials $f(x)$ and $g(x)$ are relatively prime and for some polynomial $h(x)$, $f(x)|h(x)$ and $g(x)|h(x)$, show that $f(x)g(x)|h(x)$.

9. Given non-zero polynomials $f(x)$, $g(x)$, and $h(x)$. If $f(x)$ and $g(x)$ are relatively prime, $f(x)$ and $h(x)$ are also relatively prime, show that $\text{HCF}\,(f(x), g(x)h(x)) = 1$.

10. Let $f(x)$ and $g(x)$ be non-zero polynomials and $h(x)$ be any polynomial. If $\text{HCF}\,(f(x), g(x)) = 1$, then prove that

$$\text{HCF}\,(f(x), g(x)h(x)) = \text{HCF}\,(f(x), h(x)) \ .$$

Is the converse true?

11. For any non-zero polynomials $f(x)$ and $g(x)$, prove that

$$\text{HCF}\left(\frac{f(x)}{\text{HCF}(f(x), g(x))}, \frac{g(x))}{\text{HCF}(f(x), g(x))}\right) = 1 \ .$$

12. For any non-zero polynomials $f(x)$ and $g(x)$, prove that $\text{HCF}\,(f(x), g(x)) = 1$ if and only if $\text{HCF}\,(f(x)g(x), f(x) + g(x)) = 1$.

13. Prove Corollary 2.4.8 by mathematical induction.

14. Given $p_1(x)$, $p_2(x)$, $\ldots$, $p_n(x)$ are non-associate irreducible polynomials. If, for $i = 1, 2, \ldots, n$, $p_i(x)|f(x)$, show that $\left[\prod_{i=1}^{n} p_i(x)\right]|f(x)$.

15. Let $f(x)$ and $g(x)$ be relatively prime non-constant polynomials of degrees $n$ and $m$ respectively.

    (a) Show that there exist polynomials $a_0(x)$ of degree at most $m-1$ and $b_0(x)$ of degree at most $n-1$ such that

    $$a_0(x)f(x) + b_0(x)g(x) = 1 \ .$$

    (b) Show that every pair of polynomials $a(x)$ and $b(x)$ satisfying

    $$a(x)f(x) + b(x)g(x) = 1$$

    has the form $a(x) = a_0(x) + c(x)g(x)$, and $b(x) = b_0(x) - c(x)f(x)$, for some polynomial $c(x)$.

16. Given $f(x)$ is an irreducible polynomial and $c$ is a root of $f(x)$. If $g(x)$ is a polynomial such that $g(c) \neq 0$, show that there exists a polynomial $b(x)$ of degree less than that of $f(x)$ and $g(c)b(c) = 1$.

17. Given $f(x)$ and $g(x)$ are non-constant polynomials of degree $n$ and $m$ respectively. Show that there exist non-zero polynomials $a(x)$, $b(x)$ of degree at most $m-1$ and $n-1$ respectively such that $a(x)f(x) + b(x)g(x) = 0$ if and only if $f(x)$ and $g(x)$ are not relatively prime.

18. If $f(x)$ and $g(x)$ are non-constant polynomials such that $f(x)|(g(x))^n$ for some positive integer $n$, show that either $f(x)|g(x)$ or $f(x)$ is reducible.

19. If $f(x)$, $g(x)$, and $q(x)$ are non-zero polynomials such that $q(x)(f(x))^2 = (g(x))^2$, show that $f(x)|g(x)$.

20. $d(x)$ is an HCF of $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$ if the following conditions are satisfied:
    (i) $d(x)|f_i(x)$ for $i = 1, 2, \ldots, n$,
    (ii) if $d'(x)|f_i(x)$ for each of the $f_i(x)$, then $d'(x)|d(x)$.
    Now, if $d_n(x)$ is an HCF of $f_1(x)$, $\ldots$, $f_n(x)$ and $d(x)$ is an HCF of $d_n(x)$ and $f_{n+1}(x)$, show that $d(x)$ is the HCF of $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$ and $f_{n+1}(x)$.

21. By using Question 20, find the HCF of the following polynomials
$$f_1(x) = x^3 - 6x^2 + 11x - 6 \ ,$$
$$f_2(x) = x^3 - 4x^2 + 5x - 2 \ ,$$
and
$$f_3(x) = x^3 - 5x^2 + 7x - 3 \ .$$

22. If $d(x)$ is the HCF of $a(x)$, $b(x)$, and $c(x)$, show that there exist polynomials $p(x)$, $q(x)$, and $r(x)$ such that $d(x) = a(x)p(x) + b(x)q(x) + c(x)r(x)$.

23. Let $f_1(x), f_2(x), \ldots, f_n(x)$ be non-zero polynomials and $A = \{a_1(x)f_1(x) + a_2(x)f_2(x) + \cdots + a_n(x)f_n(x) : a_i(x) \in \mathbf{F}[x]\}$.
    (a) Show that $A$ is non-empty.
    (b) Show that if $s(x)|f_i(x)$ for $i = 1, 2, \ldots, n$, then $s(x)|p(x)$ for any $p(x)$ of $A$.
    (c) Suppose $d(x)$ is a polynomial of the least degree in $A$. Show that $d(x)$ is an HCF of $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$.

## 2.5 Unique factorization theorem

We have now seen that the theory of divisibility for $\mathbf{Z}$ is entirely similar to that for $\mathbf{R}[x]$ . To conclude this chapter we shall state and prove the unique factorization theorem for polynomials of $\mathbf{R}[x]$.

2.5.1 UNIQUE FACTORIZATION THEOREM. *Every non-constant polynomial $f(x)$ of $\mathbf{R}[x]$ can be written as a product of irreducible polynomials of $\mathbf{R}[x]$. Moreover if*

$$f(x) = p_1(x) \cdots p_r(x) = q_1(x) \cdots q_s(s)$$

*where $p_i(x)$ and $q_i(x)$ are all irreducible, then $r = s$ and the order of the factors can be so arranged that each $p_i(x)$ is associated to $q_i(x)$.*

PROOF: We shall first prove that every polynomial is a product of irreducible factors. Suppose this were not true. Then the set $S$ of all polynomials that fail to be such a product would be non-empty. Select any $f(x) \in S$ with the lowest degree, i.e. $\deg f(x) \leq \deg g(x)$ for all $g(x) \in S$. Then $f(x)$ cannot be irreducible otherwise $f(x) = f(x)$ would a representation of $f(x)$ as a product of one irreducible polynomial. Thus $f(x)$ is reducible; we may write $f(x) = g_1(x)g_2(x)$ as a product of non-units. Thus, $1 \leq \deg g_1(x) < \deg f(x)$ and $1 \leq \deg g_2(x) < \deg f(x)$. If both $g_1(x)$ and $g_2(x)$ are products of irreducible factors, then $f(x)$ would be such a product which is impossible. Therefore at least one of them, say $g_1(x)$ must fail to be such a product. But this would mean that $g_1(x) \in S$ and $\deg g_1(x) < \deg f(x)$, contradicting the definition of $f(x)$ as an element of $S$ with the lowest degree. Therefore the assumption that $S$ is non-empty must be rejected. Thus every polynomial is a product of irreducible factors.

The second statement of the theorem says effectively that the irreducible factors of $f(x)$ are unique up to the order in which they appear in the product and their associates. Consider the first factor $p_1(x)$ of the first product. It follows from $p_1(x)|q_1(x) \cdots q_s(x)$ that $p_1(x)|q_j(x)$ for some factor $q_j(x)$ of the second product. By a suitable arrangement of the factors of the second product, we may assume that $p_1(x)|q_1(x)$. But both $p_1(x)$ and $q_2(x)$ are irreducible; therefore they are associates of each other. Deleting these two factors, we consider the shorter products $p_2(x) \cdots p_r(x)$ and $q_2(x) \cdots q_s(x)$. They may fail to be equal but clearly remain associates of

each other. Applying the same argument to $p_2(x)$, we find that $p_2(x)$ and $q_2(x)$ are associates after some suitable arrangement. This process can be carried on until each $p_i(x)$ is paired off with the corresponding $q_i(x)$ as associates. At this stage with all associated factors deleted, we are left with 1 in the first product and $q_{r+1}(x)\cdots q_s(x)$ in the second product. These being associates, we conclude that $r = s$. The proof of the theorem is now complete.

2.5.2 REMARKS. By the unique factorization theorem, irreducible polynomials may be taken as the basic building blocks from which all polynomials can be put together by multiplication. A domain in which the unique factorization theorem holds is called a *unique factorization domain* (UFD for short). Thus $\mathbf{R}[x]$ is a UFD. Clearly $\mathbf{Q}[x]$ and $\mathbf{C}[x]$ are also UFDs as all theorems of divisibility that we have proved for $\mathbf{R}[x]$ hold for $\mathbf{Q}[x]$ and $\mathbf{C}[x]$. We have seen in Section 2.2 that in $\mathbf{Q}[x]$ there are irreducible polynomials of every degree. Therefore after an application of the unique factorization on a polynomial of $\mathbf{Q}[x]$, no general statement on the degrees of the irreducible factors can be made. On the other hand the irreducible polynomials of $\mathbf{C}[x]$ are the linear polynomials while the irreducible polynomials $\mathbf{R}[x]$ are the linear polynomials and the quadratic polynomials with negative discriminant. Therefore every polynomial can be factorized as a product of linear polynomials in $\mathbf{C}[x]$ and every polynomial with real coefficients can be factorized as a product of linear polynomials and irreducible quadratic polynomials of $\mathbf{R}[x]$.

As for polynomials with integer coefficients, though some of the theorems in this chapter may not hold for the domain $\mathbf{Z}[x]$ and some others can only be proved quite differently, it is nevertheless true that $\mathbf{Z}[x]$ is also a UFD. However the proof of this fact is difficult and cannot be obtained by simple modification. Finally we also take note that the unique factorization theorem holds for the polynomial domains $\mathbf{Z}[x_1, \ldots, x_n]$, $\mathbf{Q}[x_1, \ldots, x_n]$, $\mathbf{R}[x_1, \ldots, x_n]$ and $\mathbf{C}[x_1, \ldots, x_n]$. Thus these are all unique factorization domains for any number $n$ of indeterminates.

CHAPTER THREE

# NOTES ON THE STUDY OF EQUATIONS IN ANCIENT CIVILIZATIONS

Equations are among the topics of mathematics that have been studied extensively for thousands of years. As equations will be the main subject for the rest of the present course, we shall begin here with a brief description of a small selection of results obtained by mathematicians in the antiquity.

## 3.1 Ancient Egyptian and Babylonian algebra

In the nineteenth century archaeologists found very old Egyptian manuscripts at burial sites in the Nile valley. These manuscripts were written in ink on a kind of paper made from the *papyrus* plants. Among these ancient manuscripts there were books on mathematics. Of these early books on mathematics the most famous is probably the Rhind papyrus now kept in the British Museum. The Rhind papyrus was written some time between 2000 B.C. and 1800 B.C. and contains numerous mathematical problems of the day; they are presented in the form of teacher's questions and pupil's answers. A very large part of this oldest surviving mathematics textbook of the world consists of practical problems of the daily life similar in mathematical content to the present-day primary school arithmetic. But there are also problems that could very well belong to secondary school algebra. These problems do not concern specific concrete objects such as bread and beer, nor are they exercises of operation on known numbers. They are actually problems on equations. The unknown ($x$ in our notation) is usually called *aha* that means heap. Problem 24 of the Rhind papyrus is an example of the aha calculation. It asks the *value of heap if heap and a seventh of heap is 19*. Written in our notation, it is to solve for $x$ in the equation

$$x + \frac{1}{7}x = 19 .$$

The Egyptian way of solving this linear equation by the *method of false position* proceeds as follows. If the value of heap is 7 then heap and a seventh heap is 8. Now 8 multiplied by $2 + \frac{1}{4} + \frac{1}{8}$ (this is the Egyptian way of writing the faction $\frac{19}{8}$) is 19. Therefore the correct value of heap is 7 multiplied by $2 + \frac{1}{4} + \frac{1}{8}$ which is $16 + \frac{1}{2} + \frac{1}{8}$ $(= \frac{133}{8})$. The solution may look extremely cumbersome today, however if we were only allowed the use of fractions with numerator 1 we would not be able to do better.

While there is no material support to think that the Egyptians knew much about algebra beyond linear equations, the ancient Babylonians were accomplished algebraists. The Egyptian way of writing is very much like our own except that the ink and the paper were different from ours; the Babylonians 'wrote' differently. They 'wrote' on clay. Wedge-shaped marks were impressed with a stylus upon soft clay tablets which were then baked hard in an oven or by the heat of the sun. This type of writing is known as *cuneiform* because of the shape of individual impressions. Clay tablets survived much better than papyrus manuscripts and thousands were found by archaeologists in the last two hundred years, now preserved in museums. From this material historians are able to study the civilization of Mesopotamia between 1500 B.C. and 1000 B.C.. Many of these tablets were identified as mathematical tables and texts.

Besides being able to solve linear equations, the Babylonians were also proficient in coping with quadratic equations and various systems of equations. For example, one of the tablets contains the following problem. *To find the side of a square if the area less than the side is the given number* $14 \times 60 + 30$ (this is the way in which numbers are written in the ancient *hexagesimal* numeral system in which the place values are powers of 60 instead of being powers of 10 as in our decimal system). In modern notation this is a quadratic equation of the form

$$x^2 - px = q \quad \text{with positive } p \text{ and } q .$$

For this type of equation, the solution given in the tablet is $x = \sqrt{(\frac{p}{2})^2 + q} + \frac{p}{2}$ .

The ancient Babylonians also studied the general solution of quadratic equations of the form

$$x^2 + bx = q$$
$$x^2 - q = px$$

where $p$ and $q$ were positive numbers. Naturally the equation $x^2 + px + q = 0$ was omitted because it may have no positive root for some positive $p$ and $q$ and negative numbers were not known to the Babylonians.

Many tables containing squares, cubes, square roots and cube roots of numbers in hexagesimal numerals were found among the clay tablets. With these tables, the Babylonians were able to find every accurate numerical solutions of equations. The Babylonians were truely the most accomplished algebraists of the ancient world.

The reader may now ask, how was it possible for the historians of today to understand the content of these ancients texts which were written in languages that have been dead for thousands of years? What was the clue? The first answer to this mystery is that about 200 years ago a stone was found in Rosetta on the west bank of the Nile – the famous Rosetta stone now kept in the British Museum. A text is carved on the Rosetta stone in Greek and two scripts of the ancient Egyptian language. Using these parallel texts as a kind of dictionary, linguists are able to decipher ancient Egyptian manuscripts.

For the Babylonian language, a 'dictionary' was found in the form of a gigantic rock cliff in Behistum, Iran. On the Behistum cliff is carved a scene of King Darius' conquest over nine neighbouring kingdoms. It also has an accompanying trilingual text in the Persian, the Babylonian and another Asian languages. With the aid of this trilingual text and the knowledge of ancient Persian, linguists are able to read ancient Babylonian.

## 3.2 Ancient Chinese algebra

In the *Jiu Zhang Suen Shu* ( 九章算書 ) of the Han Dynasty (206 B.C. – A.D. 220), we find a systematic method of solving systems of linear equations which is almost identical to the modern method

of matrix transformation. Chapter 8 of ***Jiu Zhang Suen Shu*** begins with

> 方程以御錯糅正負
> 今有上禾三秉，中禾二秉，下禾一秉，實三十九斗；上禾二秉，中禾三秉，下禾一秉，實三十四斗；上禾一秉，中禾二秉，下禾三秉，實二十六斗。問上、中、下禾實一秉各幾何。答曰：上禾一秉，九斗四分斗之一，中禾一秉，四斗四分斗之一，下禾一秉二斗四分斗之三。

Translated into English the problem and its answer are as follows:

> The yield of 3 sheaves of superior grain, 2 sheaves of medium grain and 1 sheaf of inferior grain is 39 *dou*. The yield of 2 sheaves of superior grain, 3 sheaves of medium grain and 1 sheaf of inferior grain is 34 *dou*. The yield of 1 sheaf of superior grain, 2 sheaves of medium grain and 3 sheaves of inferior grain is 26 *dou*. What is the yield of 1 sheaf of each grain?
>
> Answer: 1 sheaf of superior grain $9\frac{1}{4}$ *dou*, 1 sheaf of medium grain $4\frac{1}{4}$ *dou*, 1 sheaf of inferior grain $2\frac{3}{4}$ *dou*.

In modern notation the problem is to solve for $x, y$ and $z$ in

$$3x + 2y + z = 39$$
$$2x + 3y + z = 34$$
$$x + 2y + 3x = 26$$

and the answer is $x = 9\frac{1}{4}, y = 4\frac{1}{4}, z = 2\frac{3}{4}$ .

A method of solution is given in the text as follows:

> 術曰，置上禾三秉，中禾二秉，下禾一秉，實三十九斗，于右方。中、左行列如右方。以右行上禾徧乘中行而以直徐。又乘其次，亦以直徐，然以中行中禾不盡者徧乘左行而以直除。左方下禾不盡者，上爲法，下爲實。實即下禾之實。求中禾，以法乘中行下實，而除下禾之實。餘如中禾秉數而一，即中禾之實。求上禾亦以法乘右行下實，而除下禾，中禾之實。餘如上禾秉數而一，即上禾之實，實皆如法，各得一斗。

Following the instruction, we first write down the coefficients in matrix form

$$\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{array}$$

Then we multiply the middle column throughout by the top number (superior grain) of the right column and subtract repeatedly from it the right column:

$$\begin{array}{ccc} 1 & 6 & 3 \\ 2 & 9 & 2 \\ 3 & 3 & 1 \\ 26 & 102 & 39 \end{array} \qquad \begin{array}{ccc} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{array}$$

We carry out the same operation on the left column to get

$$\begin{array}{ccc} 3 & 0 & 3 \\ 6 & 5 & 2 \\ 9 & 1 & 1 \\ 78 & 24 & 39 \end{array} \qquad \begin{array}{ccc} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{array}$$

Now we multiply the left column by the uppermost non-vanishing number (medium grain) of the middle column and subtract:

$$\begin{array}{ccc} 0 & 0 & 3 \\ 20 & 5 & 2 \\ 40 & 1 & 1 \\ 195 & 24 & 39 \end{array} \qquad \begin{array}{ccc} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{array}$$

By now we have transformed the original system into

$$\begin{aligned} 3x + 2y + z &= 39 \\ 5y + z &= 24 \\ 36z &= 99 \ . \end{aligned}$$

The rest of the instruction is just easy evaluation of the unknowns $x, y$ and $z$ in the obvious manner.

We remark that in ancient China numerical calculation was not carried out on an abacus (算盤) of the kind that is still obtainable in shops in Hong Kong but on a *counting board* with *counting rods* (算籌). On a counting board the initial matrix would look something like the figure below:

| | | |
|---|---|---|
| 𝍠 | 𝍡 | 𝍢 |
| 𝍡 | 𝍢 | 𝍡 |
| 𝍢 | 𝍠 | 𝍠 |
| 𝍪𝍥 | 𝍪𝍣 | 𝍫𝍨 |

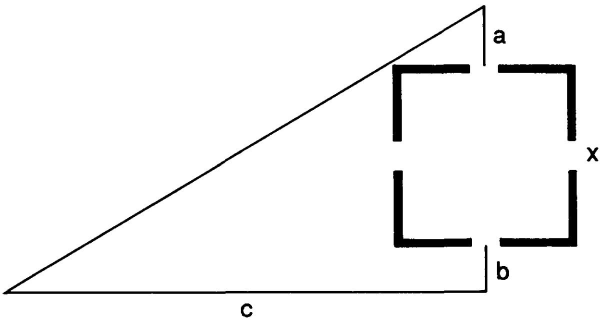
Counting rods are put in or taken from the fields of the board as the transformation is being carried out. Therefore this ancient 'calculator' is extremely well suited for the instruction of *Jiu Zhang Suen Shu*.

Quadratic equations are treated in Chapter 9 of the ancient textbook. Problem 20 reads:

> 今有邑方不知大小，各中開門。出北門二十步有木。出南門十四步，折而西行一千七百七十五步見木。問邑方幾何。
> 答曰：二百五十步。

> A square city is of side unknown *pu*. At the centre of the wall on each side is a gate. 20 *pu* from the north gate is a tree. If one comes out of the south gate, walks 14 *pu*, turns west and walks another 1775 *pu*, one would see the tree. How many *pu* is the side of the square city?
>
> Answer: 250 *pu*.

According to the text we have the following map of the city:



Therefore the problem is to solve for the unknown $x$ in the equation

$$x^2 + (a+b)x = 2ac \ .$$

Let us now read the instruction of solution given in the text.

術曰：以出北門步數乘西行步數，倍之，爲實。並出南門步數爲從法，開方除之，即邑方。

Method: Multiply the number of *pu* from the tree to the north gate by the number of *pu* of the westward walk. Double the product to form the *Shi*. Add [to the numbers of *pu* from the tree to the north gate] the number of *pu* from the south gate [to the point of turning] to form the *Cong fa*. Apply [to the *Shi* and the *Cong-fa* the method of] taking root and subtracting to obtain the answer.

The *Shi* of text refers to the constant term $2ac$ and the *Cong-fa* the linear coefficient $a+b$ of the equation. The method of taking root and subtracting is the standard routine to obtain

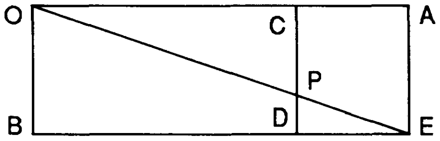
$$x = \frac{\sqrt{(a+b)^2 + 8ac} - (a+b)}{2} = 250 \ .$$

### 3.3 Ancient Greek algebra

There is enough material evidence that the ancient Greeks learnt their mathematics (particulary algebra) from the Babylonians and redeveloped it from its foundation into a glorious edifice. By the full employment of deductive reasoning the Greek philosopher–mathematicians turned the ancient empirical mathematics of the Egyptians and the Babylonians into a rigorous theoretical science. Greek geometry is no doubt a shining example of this achievement. The equally illustrious *geometric algebra* developed by the Greeks, however, has attracted less attention and admiration partly because it is entirely formulated in geometric terms and partly because it is no longer taught at schools. In actual fact, geometric algebra is an integral part of the Greek geometry — no less than three of the thirteen books of Euclid's *Elements* (300 B.C.) are devoted to geometric algebra and arithmetic.

It is intructive to see how the Greeks formulated and solved linear and quadratic equations. In the fifth century B.C. the solution of the linear equation $ax = bc$ would mean the construction of a rectangle with one side given as $a$ to have the same area as a given rectangle with sides $b$ and $c$. The construction is carried out as follows.

Draw a rectangle $OCDB$ with $b = OB$ and $c = OC$ . On $OC$ lay off $OA$ so that $OA = a$. Complete the rectangle $OAEB$ and draw the diagonal $OE$ to cut $CD$ at $P$. Then $CP = x$ is the other side of the desired rectangle. Because $\frac{CP}{AE} = \frac{OC}{OA}$ , we get $ax = bc$.



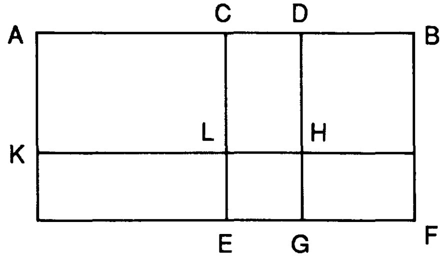

The well-known algebraic identity

$$(a+b)(a-b) = a^2 - b^2$$

is found in Proposition 5, Book II of *Elements*. The proposition is quoted below where the insertions within brackets are our elucidation of the unfamiliar formulation.

> **If a straight line [$2a$] be cut into equal [$a$ and $a$] and unequal [$a+b$ and $a-b$] segments, the rectangle [$(a+b)(a-b)$] contained by the unequal segments of the whole, together with the square [$b^2$] on the straight line between the points of section is eqal to the square [$a^2$] on the half.**

Euclid's proof is illustrated in the diagram below when $AC = CB = CE = a$ and $CD = LE = b$.



Here

$$(a+b)(a-b) = ADHK = DBFG + CDHL$$

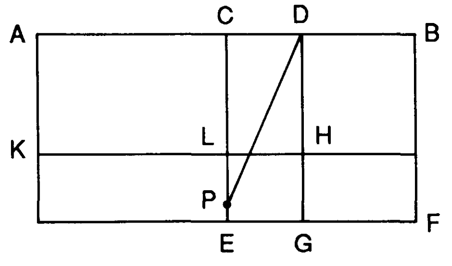
$$a^2 = CBFE = DBFG + CDHL + LHGE$$

Therefore $a^2 = (a+b)(a-b) + b^2$.

The figure above used in the construction of Proposition 5 proves to be a very valuable tool for solution of quadratic equations in the Greek fashion. Take for instance the equation

$$ax - x^2 = b^2$$

with $a > 2b$ . A student of Euclid at the University of Alexandria would begin with a line segment $AB = a$. He would then bisect $AB$ at $C$ and erect a perpendicular $CP = b$ . With a compass a point $D$ on $AB$ with $DP = \frac{a}{2}$ is then found. After completing a similar figure with the four point $A, B, C, D$, he would obtain the solution $x = DB$. Because by Proposition 5, $x(a-x)$ + area $LHGE = (\frac{a}{2})^2$ and by Pythagoras' theorem, the square on $PD(\frac{a}{2})$ is the sum of the square on $CP(b)$ and the square on $CD$ .




### 3.4 The modern notations

In the last three sections we have a very brief survey of ancient algebra from 2000 B.C. – A.D. 220. For the later development of the subject, we shall have to refer the interested reader to books on history of mathematics. To round off here, we take a quick glance at the emergence of the modern notations of algebra. This took a very long time to develop. In fact the modern way of writing equations

in the form such as $3x + 6 = 0, x^5 + 4x^3 - 3x - 8 = 0$ was not invented until the seventeenth century. In the sixteenth century François Viéte (1540–1603) wrote the equation $x^3+3B^2x = 2z^3$ in the very antiquated form

$$A \ cubus \ + \ B \ plano \ 3 \ in \ A, \ aequari \ Z \ solido \ 2 \ .$$

Thomas Harriot (1560–1621) has a better set of notations. For $52 = -3a + a^3$ he wrote

$$52 \;=\!=\!=\; -3 \cdot a + aaa$$

with an elongated equality sign. René Descartes (1596–1650) was the first to suggest the use of letters $x, y, z$ for the unknowns and he came very close to the notation of today. For example, he wrote

$$x^3 - \ -9xx + 26x - \ -24 \propto 0$$

for the equation $x^3 - 9x^2 + 26x - 24 = 0$. By late seventeenth century, European mathematicians were able to use the modern notations and carry out manipulations on symbols in much the same way as we do today.

CHAPTER FOUR

# LINEAR, QUADRATIC AND CUBIC EQUATIONS

A polynomial $g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ defines a polynomial function $g(x) : \mathbf{R} \to \mathbf{R}$ which maps every real number $c$ of the domain to the real number $g(c)$ of the range. The evaluation of $g(x)$ at $x = c$ is a very staight-forward matter and there are simple methods of calculation by which the correct value of $g(c)$ can be obtained. We are now interested in the possibility of finding real values $c$ of the domain such that $g(c)$ coincides with an pre-assigned value $d$ of the range. Thus given $g(x) \in \mathbf{R}[x]$ and $d \in \mathbf{R}$, we seek information on the possible values of $c$ such that $g(c) = d$. In the language of set theory, the problem is to find the pre-images $c$ of $d$ under the mapping $g(x) : \mathbf{R} \to \mathbf{R}$. After absorbing the number $-d$ into the constant term of $g(x)$, i.e. replacing $g(x)$ by $f(x) = g(x) - d$, this amounts to the evaluation of all real roots $c$ of the polynomial function $f(x)$ .

In contrast to the evaluation of a polynomial $f(x)$ at a given value of $x$, the problem of finding roots of a given polynomial function $f(x)$ is a very difficult problem of mathematics. In this chapter, we shall study the methods of solving some simple equations.

## 4.1 Terminology

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial in the indeterminate $x$ with real coefficients. If we regard the symbol $x$ in the above expression as a *definite* but *unknown* real or complex number, then the expression simply represents a number. Since numbers can be compared by equality, it is therefore legitimate to say that we wish

(A) To find the values of the unknown number $x$ such that $a_n x^n + a_n x^{n-1} + \cdots + a_1 x + a_0 = 0$.

This being the problem at hand, we may also say (A) in any one of the following ways:

(B) To solve for $x$ in the *polynomial equation*

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \ .$$

(C) To find all *roots* of the *equation*

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \ .$$

Furthermore given a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ in the indeterminate $x$, we may also use the abbreviated expression

$$f(x) = 0$$

for the equation

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$$

in the unknown $x$. Terminology such as *degree, coefficients, terms,* etc. of an equation in the unknown $x$ shall have the obvious meaning. Moreover a *solution*, a *root* and a *zero* of an equation $f(x) = 0$ all mean a real or complex number $c$ such that $f(c) = 0$ .

4.1.1 REMARKS. A sharp distinction must be made between the equation $f(x) = 0$ in the unknown $x$ and the equality $f(x) = 0$ of polynomials in the indeterminate $x$. In the former case, $x$ is a definite (though unknown) number and the expression $f(x)$ is also a definite number. Therefore the equation $f(x) = 0$ is to be correctly interpreted as the condition on this number $x$ that the associated number $f(x)$ should be zero. In the latter case $f(x)$ is a polynomial and so is 0; the equality of these two polynomials means that all coefficients $a_i$ of $f(x)$ are zero. Therefore the equality $f(x) = 0$ of polynomials is the condition on the coefficients $a_i$ of $f(x)$ that they should all be zero.

There are many other kinds of equations besides polynomial equations in one unknown. In the first place there are polynomial equations in two or more unknowns $x, y, \ldots$ . Then there are equations which are not polynomial equations. For example, if $f(x)$ and $g(x) \neq 0$ are polynomials in the indeterminate $x$, then the equation

$f(x)/g(x) = 0$ would be a rational equation in the unknown $x$; moreover, unless $g(x)$ is a factor of $f(x)$, it is not a polynomial equation. An expression such as $\cos^2 x + 3\sin x + 5 = 0$ would be a trigonometrical equation in the unknown $x$, and $x + 7^{x+8} = 0$ would be an exponential equation in the unknown $x$. Here we are only interested in polynomial equations with real coefficients in one unknown, their properties and their solutions.

In the subsequent sections of this chapter we shall use the results of the previous chapters to study the problem (A). To conclude the present section, we observe that a number $c$ is a root of the equation $f(x) = 0$ in the unknown $x$ if and only if the linear polynomial $x - c$ is a factor of the polynomial $f(x)$. This reformulation of the factor theorem leads to the following obvious results.

4.1.2 THEOREM. *An equation $f(x) = 0$ in the unknown $x$ of degree $n \geq 1$ has at most $n$ distinct roots.*

## 4.2 Linear and quadratic equations

For completeness and convenient reference we record here some well-known results on solution of equations in one unknown of degree less than three.

The trivial equation

$$0x = 0$$

in the unknown $x$ admits every value of $x$ as a solution. It is the only polynomial equation that has an infinite number of roots.

An equation of degree 0

$$0x + a = 0 \quad (a \neq 0)$$

in the unknown $x$ has no root.

A linear equation

$$ax + b = 0 \quad (a \neq 0)$$

in the unknown $x$ admits a unique solution and it is $-b/a$ .

A quadratic equation

$$ax^2 + bx + c = 0 \quad (a \neq 0)$$

in the unknown $x$ with real coefficients has a single real root and it is $-b/2a$ if and only if the discriminant $D = b^2 - 4ac = 0$. The equation has two distinct real roots if and only if $D > 0$. In this case the roots are $(-b \pm \sqrt{D})/2a$. The equation has two distinct imaginary roots if and only if $D < 0$ . In this case the roots are the complex conjugates $(-b + i\sqrt{-D})/2a$ and $(-b - i\sqrt{-D})/2a$ .

Conversely it follows from the Factor Theorem 1.5.2 that given any two numbers $\alpha$ and $\beta$,

$$x^2 - (\alpha + \beta)x + \alpha\beta = 0$$

is a quadratic equation whose roots are exactly $\alpha$ and $\beta$.

## EXERCISE 4A

1. Solve the equation $3x^2 - 2x + k = 0$ for real number $k$.
2. Find the values of $m$ such that the real quadratic equation

$$(m - 1)x^2 + 2mx + m + 3 = 0$$

has real roots and solve the equation for these values of $m$.

3. If $a$, $b$, and $c$ are real numbers such that $3a$, $b$ and $2c$ are in A.P., prove that the equation $ax^2 + bx + c = 0$ has real roots.
4. If $m > n > 0$, prove that the equation $2x^2 - (3m + n)x + mn = 0$ has two unequal real roots, one is greater than $n$ and the other is less than $n$.
5. Consider $x^2 + (3 + 4i)x - (14 - 6i) = 0$. Find its discriminant $D$ and show that $D > 0$ but the equation has two complex roots. Thus the discriminant test for quadratic equations of real coefficients fails for complex coefficients.

Numbers 6 to 9 give a series of properties of roots of quadratic equations of complex coefficients, in which $a$, $b$, $c$, and $d$ are real numbers. Prove these properties.

6. The quadratic equation

$$x^2 + (a + bi)x + c + di = 0$$

has two unequal real roots if and only if

$$\begin{cases} b = d = 0 \,, \\ a^2 - 4c > 0 \,. \end{cases}$$

How about if the equation has equal real roots?

7. The quadratic equation

$$x^2 + (a + bi)x + c + di = 0$$

has two conjugate complex roots if and only if

$$\begin{cases} b = d = 0 \,, \\ a^2 - 4c < 0 \,. \end{cases}$$

8. The quadratic equation

$$x^2 + (a + bi)x + c + di = 0$$

has only one real root and the other root is complex if and only if

$$\begin{cases} b \neq 0 \,, \\ d^2 - abd + b^2c = 0 \,. \end{cases}$$

9. The quadratic equation

$$x^2 + (a + bi)x + c + di = 0$$

has two non-conjugate complex roots if and only if

$$\begin{cases} b \neq 0 \,, \\ d^2 - abd + b^2c \neq 0 \end{cases} \quad \text{or} \quad \begin{cases} b = 0 \,, \\ d \neq 0 \,. \end{cases}$$

## 4.3 Cubic equations

The Babylonians knew how to solve quadratic equations almost four thousand years ago. It was more than three thousand years later that a general method for solving cubic equations was available. The method was discovered in Italy by Scipione del Ferro (*ca.* 1465–1526) and Nicolo Tartaglia (*ca.* 1500–1557) and was made known in the

*Ars magna* by Geromino Cardano (1501–1576). The method consists of successive reductions of a given cubic equation to more convenient forms.

Let

$$x^3 + ax^2 + bx + c = 0$$

be a cubic equation with real coefficients. The first reduction for the sole purpose of eliminating the quadratic term is carried out by replacing the unknown $x$ by another unknown quantity $y - \frac{1}{3}a$. The original equation in $x$ is then transformed into the *intermediate equation*

$$y^3 + px + q = 0$$

in a new unknown $y$ with a vanishing quadratic term. The coefficients of the intermediate equation is easily seen to be

$$p = b - \frac{1}{3}a^2, \quad q = c - \frac{1}{3}ab + \frac{2}{27}a^3 .$$

Every root of the intermediate equation in the unknown $y$ will give rise to a root of the original equation in the unknown $x = y - \frac{1}{3}a$. The second reduction aims at eliminating the linear term of the intermediate equation. For this purpose we replace in the intermediate equation $y^3 + px + q = 0$ the unknown $y$ by another quantity $z_1 + z_2$, and obtain

$$z_1{}^3 + z_2{}^3 + (3z_1z_2 + p)(z_1 + z_2) + q = 0 .$$

As this equation does not have vanishing linear terms, we shall impose on the unknown quantities $z_1$ and $z_2$ the extra condition

$$3z_1z_2 + p = 0 .$$

Therefore the original problem is now reduced to finding the unknown values of $z_1$ and $z_2$ such that

$$\begin{cases} z_1{}^3 + z_2{}^3 = -q \\ z_1z_2 = -\frac{1}{3}p . \end{cases}$$

The sum of each pair of such values $z_1$ and $z_2$ will be a root $y = z_1 + z_2$ of the intermediate equation $y^3 + px + q = 0$ which in turn will give us a root $x = y - \frac{1}{3}a$ of the original cubic equation $x^3 + ax^2 + bx + c = 0$ .

Unfortunately it is not easy to solve this set of simultaneous equations in $z_1$ and $z_2$ directly. The interested reader will find that an application of the usual method of elimination to these equations would lead to an equation of degree 6. To overcome this difficulty, we consider a second set of simultaneous equations

$$\begin{cases} {z_1}^3 + {z_2}^3 = -q \\ {z_1}^3 {z_2}^3 = -\frac{1}{27}p^3 \end{cases} .$$

Clearly every pair of solutions to the first set of equations is a pair of solutions to the second set of equations. But the converse may not be true.

Now this second set of equations can be viewed as a condition on the numbers ${z_1}^3$ and ${z_2}^3$. Adopting this point of view we shall be looking for two quantities ${z_1}^3$ and ${z_2}^3$ whose sum is $-q$ and whose product is $-\frac{1}{27}p^3$. But such numbers ${z_1}^3$ and ${z_2}^3$ are precisely the roots of the quadratic equation

$$(z^3)^2 + qz^3 - \frac{1}{27}p^3 = 0$$

in the unknown $z^3$ .

Now this final quadratic equation in $z^3$ is easy to solve and the values of ${z_1}^3$ and ${z_2}^3$ are simply

$${z_1}^3 = -\frac{q}{2} + \sqrt{r} \quad \text{and} \quad {z_2}^3 = -\frac{q}{2} - \sqrt{r}$$

where $r = \frac{1}{4}q^2 + \frac{1}{27}p^3$ .

Taking cube roots of these two numbers will give rise to 3 values for each $z_1$ and $z_2$ and hence to 9 pairs of solutions of the second set of equations. Among them we are only interested in those that are solutions of the first set of equations, i.e. those pairs of $z_1$ and $z_2$ such that

$$z_1 z_2 = -\frac{1}{3}p .$$

Let us proceed to find these pairs. We start off by picking any $z_1$ and $z_3$ such that

$${z_1}^3 {z_2}^3 = -\frac{1}{27}p^3 .$$

Then for their product it is either

$$z_1 z_2 = -\frac{1}{3}p, \quad z_1 z_2 = -\frac{1}{3}p\omega \text{ or } z_1 z_2 = -\frac{1}{3}p\omega^2$$

where $\omega = -\frac{1}{2}(1 - i\sqrt{3})$ is a primitive cube root of unity. If it is the first case, then the pair $z_1$ and $z_2$ have the required property. In the second case, the pair $\omega z_1$ and $\omega z_2$ of cube roots of $z_1{}^3$ and $z_2{}^3$ will do. In the last case, the pair $\omega z_1$ and $z_2$ will do. Therefore among the cube roots of $z_1{}^3$ and $z_2{}^3$ we have pairs that are solutions to the first set of equations.

Let us take any one such pair and denote them by

$$s = \sqrt[3]{-\frac{1}{2}q + \sqrt{r}} \quad \text{and} \quad t = \sqrt[3]{-\frac{1}{2}q - \sqrt{r}}\,,$$

thus $3st = -p$ . Then $s + t$ is a root of the intermediate equation $y^3 + px + q = 0$ (as we may verify by direct substitution) and $s + t - \frac{1}{3}a$ is a root of the original equation $x^3 + ax^2 + bx + c = 0$ . The other roots of the intermediate equation are $\omega s + \omega^2 t$ and $\omega^2 s + \omega t$ . To see this, we merely have to verify the following equality of polynomials

$$(y - (s + t))(y - (\omega s + \omega^2 t))(y - (\omega^2 s + \omega t)) = y^3 + py + q\ .$$

Aternatively we may argue as follows. The three possible values for $z_1$ are $s$, $\omega s$, $\omega^2 s$ and those for $z_2$ are $t$, $\omega t$, $\omega^2 t$. But only the three pairings $s$ with $t$, $\omega s$ with $\omega^2 t$ and $\omega^2 s$ with $\omega t$ will satisfy the requirement that $3z_1 z_2 = -p$. Hence the roots of $y^3 + py + q = 0$ are

$$s + t, \quad \omega s + \omega^2 t \quad \text{and} \quad \omega^2 s + \omega t\ .$$

We summarize the above discussion in the following theorems.

4.3.1 THEOREM. *The roots of a cubic equation*

$$x^3 + ax^2 + bx + c = 0$$

*is given by Cardano's formulae as:*

$$s + t - \frac{1}{3}a, \quad \omega s + \omega^2 t - \frac{1}{3}a, \quad \omega^2 s + \omega t - \frac{1}{3}a$$

*where*

$$s = \sqrt[3]{-\frac{1}{2}q + \sqrt{r}}, \quad t = \sqrt[3]{-\frac{1}{2}q - \sqrt{r}}, \quad \text{such that} \quad 3st = -p$$

$$r = \frac{1}{4}q^2 + \frac{1}{27}p^3, \qquad p = b - \frac{1}{3}a^2, \qquad q = c - \frac{1}{3}ab + \frac{2}{27}a^3$$

and $\omega = -\frac{1}{2}(1 - i\sqrt{3})$ .

4.3.2 THEOREM. *The roots of a cubic equation*

$$y^3 + py + q = 0$$

*with vanishing quadratic terms are* $s + t$, $\omega s + \omega^2 t$ *and* $\omega^2 s + \omega t$ *where* $s, t$ *and* $\omega$ *have the values given in Theorem 4.3.1.*

4.3.3 EXAMPLE. *Solve the equation*

$$x^3 - 15x - 126 = 0 .$$

SOLUTION: This is a cubic equation with vanishing quadratic term. The solutions are given in the above theorem and the values of $s, t$ can be obtained by substitution into Cardano's formulae. The reader will agree that complicated formulae such as Cardano's are too difficult to learn by heart and that a solution by retracing the steps of the theorem would be preferred. Substitute $z_1 + z_2$ for $x$ into the given equation to get

$${z_1}^3 + {z_2}^3 + (3z_1z_2 - 15)(z_1 + z_2) - 126 = 0 .$$

The elimination of the linear coefficients leads to the conditions

$$\begin{cases} {z_1}^3 + {z_2}^3 = 126 \\ z_1z_2 = 5 . \end{cases}$$

Therefore ${z_1}^3$ and ${z_2}^3$ are the roots of the quadratic equation

$$(z^3)^2 - 126z^3 + 125 = 0$$

in the unknown $z^3$ . They are 1 and 125. Taking the cube roots we obtain for $z_1$ the values 1, $\omega$, $\omega^2$ and for $z_2$ the values 5, $5\omega$, $5\omega^2$. Now any pair

of these values will satisfy the condition $z_1{}^3 + z_2{}^3 = 126$ but not every pair will satisfy the second condition $z_1 z_2 = 5$. This restriction gives the final selection of

$$1+5, \quad \omega + 5\omega^2 \quad \text{and} \quad \omega^2 + 5\omega \ .$$

Therefore the roots of the given cubic equation are 6, $-3 - 2i\sqrt{3}$ and $-3 + 2i\sqrt{3}$.

4.3.4 EXAMPLE. *Solve the equation*

$$x^3 + 6x^2 + 3x + 18 = 0 \ .$$

SOLUTION: To eliminate the quadratic term we subsitute $y - 2$ for $x$ into the equation to get the intermediate equation

$$y^3 - 9y + 28 = 0$$

in the general form of

$$y^3 + py + q = 0$$

with $p = -9$ and $q = 28$. Instead of working with the substitution $y = z_1 + z_2$ to obtain the final quadratic equation in the unknown $z^3$ as we have done in the last example, we may also use the alternative substitution

$$y = z - \frac{p}{3z}$$

which is suggested by $y = z_1 + z_2$ and $3z_1 z_2 = -p$. Thus we replace $y$ by

$$z + \frac{3}{z}$$

in the intermediate cubic equation and multiply afterward throughout by $z^3$:

$$z^3\{(z + \frac{3}{z})^3 - 9(z + \frac{3}{z}) + 28\} = 0$$

to obtain the quadratic equation

$$(z^3)^2 + 28z^3 + 27 = 0,$$

in the unknown $z^3$. The solutions of this equation are $-1$ and $-27$. Taking cube roots we obtain

$$-1, \ -\omega, \ -\omega^2 1 \quad \text{and} \quad -3, \ -3\omega, \quad -3\omega^2 \ .$$

Therefore the roots of the intermediate cubic equation $y^3 - 9y + 28 = 0$ are

$$-4, \quad -\omega - 3\omega^2 \quad \text{and} \quad -\omega^2 - 3\omega \; .$$

Finally, subtracting 2 from each, we obtain the roots of the given cubic equation $x^3 + 6x^2 + 3x + 18 = 0$ as $-6$, $i\sqrt{3}$ and $-i\sqrt{3}$.

4.3.5 EXAMPLE. *Solve the equation*

$$x^3 - 3x^2 - 9x + 27 = 0 \; .$$

SOLUTION: The first reduction calls for the replacement of $x$ by $y+1$. Thus the intermediate equation $y^3 + px + q = 0$ is

$$y^3 - 12y + 16 = 0 \; .$$

Using the method of the last example, we substitute $z + \frac{4}{z}$ for $y$ and multiply by $z^3$ to get the final quadratic equation in $z^3$ as

$$(z^3)^2 + 16z^3 + 64 = 0 \; .$$

This has a double root $-8$. Therefore we obtain $s = t = -2$, and the roots of the intermediate equations as $-4$, $2$, $2$. Hence the roots of the given cubic equation are a simple root $-3$ and a double root $3$.

Similar to the quadratic equations, a cubic equation also has a discriminant from which full information on the nature of its roots can be obtained. Let us consider the case where the given cubic equation

$$g(y) = y^3 + py + q = 0$$

has a vanishing quadratic term. Following the procedure given earlier, we obtain a quadratic equation

$$h(z^3) = z^6 + qz^3 - \frac{p^3}{27} = 0$$

in the unknown $z^3$. Denote its discriminant by $D$ and consider $\triangle = -27D$. Then for the roots of the quadratic equation $h(z^3) = 0$ we can distinguish three different cases.

Case 1. If $\triangle < 0$ then there are two distinct real roots

$$-\frac{q}{2}+\frac{\sqrt{D}}{2} \quad \text{and} \quad -\frac{q}{2}-\frac{\sqrt{D}}{2}\,.$$

Case 2. If $\triangle = 0$ then there are two equal real roots

$$-\frac{q}{2}\,.$$

Case 3. If $\triangle > 0$ then there are two conjugate imaginary roots

$$-\frac{q}{2}+\frac{\sqrt{-D}}{2}i \quad \text{and} \quad -\frac{q}{2}-\frac{\sqrt{-D}}{2}i\,.$$

We shall study the roots of the cubic equation $g(y) = 0$ case by case according to this classification.

In Case 1, we may take for values of $s$ and $t$ in Cardano's Formulae the real cubic roots

$$s = \sqrt[3]{-\frac{q}{2}+\frac{\sqrt{D}}{2}} \quad \text{and} \quad t = \sqrt[3]{-\frac{q}{2}-\frac{\sqrt{D}}{2}}\,.$$

Then $st = -p/3$. Therefore $g(y) = 0$ has one real root $s + t$ and two imaginary roots $\omega s + \omega^2 t$ and $\omega^2 s + \omega t$.

In Case 2, we may also use the real cubic root

$$s = t = \sqrt[3]{-\frac{q}{2}}\,.$$

Again $st = -p/3$. Now it follows from $\omega^2 + \omega + 1 = 0$ that $\omega s + \omega^2 t = \omega^2 s + \omega t = -s$. Therefore $g(y) = 0$ has three real roots $2s, -s, -s$, at least two of which are equal.

Finally in Case 3, let $s = \alpha + \beta i$ be a cube root of $(-q + i\sqrt{-D})/2$. Then $t = \alpha - \beta i$ is easily seen to be a cube root of $(-q - i\sqrt{-D})/2$. Since $st$ is real *it* equals $-p/3$. Therefore $s + t = 2\alpha$, $\omega s + \omega^2 t = -\alpha - \beta\sqrt{3}$ and $\omega^2 s + \omega t = -\alpha + \beta\sqrt{3}$ are the three distinct real roots of $g(y) = 0$.

This completes the study of the roots of the equation $g(x) = 0$. Now the roots of a general cubic equation

$$f(x) = x^3 + ax^2 + bx + c = 0$$

differ from those of $g(y) = 0$ only by a real constant $a/3$. Therefore we should obtain the same results using

$$\begin{aligned}\Delta &= -4p^3 - 27q^2 \\ &= -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2 \ .\end{aligned}$$

Finally let us summarize the whole discussion above by the following statements.

4.3.6 DEFINITION. The discriminant $\Delta$ of a cubic equation $x^3+px+q=0$ is the real constant

$$\Delta = -4p^3 - 27q^2 \ .$$

4.3.7 DEFINITION. The discriminant $\Delta$ of a cubic equation $x^3 + ax^2 + bx + c = 0$ is the real constant

$$\Delta = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2 \ .$$

4.3.8 THEOREM. *Let $\Delta$ be the discriminant of a cubic equation $x^3 + ax^2 + bx + c = 0$. If $\Delta < 0$ then the equation has one real root and two distinct imaginary roots. If $\Delta = 0$ then the equation has three real roots, at least two of which are equal. If $\Delta > 0$ then the equation has three distinct real roots.*

4.3.9 REMARK. We observe that in Examples 4.3.3 and 4.3.4 we have $\Delta < 0$ and in Example 4.3.5 we have $\Delta = 0$. None of these examples presents any difficulty. In the case where $\Delta > 0$, we have to work with $D < 0$. This means the roots $(-q + i\sqrt{-D})/2$ and $(-q - i\sqrt{-D})/2$ of the final quadratic equation are both imaginary. Thus we shall have to use De Moivre's theorem to work out their cube roots. Therefore in general, if we know that the given cubic equation has three distinct real roots or has a positive discriminant, then it is not advisable to use Cardano's formulae.

## EXERCISE 4B

1. For the equation $x^3 + ax^2 + bx + c = 0$, by putting $y = x + h$, obtain an equation in $y$. Choose a value for $h$ such that the new equation is of the form $y^3 + py + q = 0$, hence find $p$ and $q$ in terms of $a$, $b$, and $c$.

2. Solve the following equations.
   (a) $x^3 - 9x^2 + 26x - 24 = 0$
   (b) $x^3 - 3x^2 + 3x - 2 = 0$
   (c) $x^3 - 6x^2 + 9x - 4 = 0$.
3. For the equation $y^3 + py + q = 0$, by putting $y = z - \frac{p}{3z}$ and $u = z^3$, obtain a quadratic equation in $u$. Hence solve the following equations.
   (a) $x^3 + 18x - 19 = 0$
   (b) $x^3 + 18x + 215 = 0$.
4. For each of the following equations, calculate the discriminant and determine the nature of the roots.
   (a) $x^3 - 4x^2 - 3x + 12 = 0$
   (b) $2x^3 + 7x^2 + 22x - 13 = 0$
   (c) $x^3 - x^2 - 8x + 12 = 0$
   (d) $x^3 - 12x^2 + 45x - 50 = 0$.
5. Let $\alpha$, $\beta$, and $\gamma$ be the roots of $x^3 + px + q = 0$.
   (a) Show that $(\beta - \gamma)^2 = -4p - 3\alpha^2$.
   (b) Express $(\beta - \gamma)^2(\gamma - \alpha)^2(\alpha - \beta)^2$ in terms of $p$ and $q$.
   (c) Show that the equation has a multiple root if and only if $4p^3 + 27q^2 = 0$.
6. If $\alpha$, $\beta$, and $\gamma$ are the roots of the equation $x^3 + ax^2 + bx + c = 0$, by using Cardano's formulae, show that the discriminant $\triangle = (\alpha - \beta)^2 \cdot (\beta - \gamma)^2 \cdot (\gamma - \alpha)^2$. Hence, determine the nature of the roots when $\triangle$ is less than, equal to or greater than zero.
7. Find the range of the real number $p$ for which the equation $2x^3 + 9x^2 + 12x + p = 0$ has three distinct real roots.
8. Determine the nature of the roots of the equation
   $$x^3 - 3x^2 + 2ax - 1 = 0$$
   for different real values of $a$.
9. If $r \neq s$, express $r$ and $s$ in terms of $p$ and $q$ such that $x^3 + px + q = \dfrac{r(x-s)^3 - s(x-r)^3}{r-s}$. Hence solve the following equations:
   (a) $x^3 - 6x + 9 = 0$
   (b) $x^3 - 9x + 28 = 0$.

10. Using the identity $\cos 3\phi = 4\cos^3\phi - 3\cos\phi$, find, in terms of $k$ and $\phi$, the roots of the equation $x^3 - \dfrac{3k^2}{4}x - \dfrac{k^3\cos 3\phi}{4} = 0$, where $k > 0$. Hence solve the following equations:

    (a) $x^3 - \frac{9}{4}x + \frac{8}{9} = 0$

    (b) $x^3 - 6x - 4 = 0$.

## 4.4 Equations of higher degree

A general method for solving the general quartic equation

$$x^4 + ax^3 + bx^2 + cx + d = 0$$

is also found in Cardano's *Ars Magna*. This method is attributed to Cardano's assistant Ludovico Ferrari (1522–1565). As in the cases of quadratics and cubics the solutions of a quartic equation are given in terms of root extractions and rational operations (i.e. addition, subtraction, multiplication and division) performed on the coefficients of the given equation. From the sixteenth century to the beginning of nineteenth century many mathematicians tried to obtain similar results for quintic equations but without complete success. It was then suspected that equations of higher degrees could not be solved by root extraction and rational operations on coefficients. This was confirmed by Paolo Ruffini (1765–1833) and Niels Hendrik Abel (1802–1829) that there is no general formula of such form. The definitive answer of this kind of study was obtained by Evariste Galois (1811–1832) who not only confirmed the results of Ruffini and Abel but also provided criteria for solvability of any $n$-th degree equation by rational operations and root extraction on coefficients. The search for general solution of equations that began with the Egyptians and the Babylonians ended with the discovery of Galois. The method that he used is now called Galois theory and is included in many standard undergraduate courses on abstract algebra. Other classical problems such as the trisection of an angle and the quadrature of a circle by ruler and compass which have exercised the best brains of the world for centuries, since they were put forward in ancient Greece, also obtain definitive answers in Galois theory.

CHAPTER FIVE

# ROOTS AND COEFFICIENTS

We remarked in the last chapter that for an equation of degree higher than four we do not possess a general method of solution and that the roots of such equations may not be obtained by root extractions and rational operations on the coefficients. Naturally this does not mean that we shall henceforth neglect the study of equations of higher degrees, but rather that we should learn individual methods to suit individual types of equations. In this chapter we pay special attention to the formal relations between the roots and the coefficients, and develop some purely algebraic methods.

## 5.1 Basic relations

We recall that the roots of a quadratic equation with leading coefficient 1

$$x^2 + b_1x + b_0 = 0$$

are given as

$$r_1 = -\frac{b_1}{2} + \frac{\sqrt{b_1{}^2 - 4b_0}}{2} \quad \text{and} \quad r_2 = -\frac{b_1}{2} - \frac{\sqrt{b_1{}^2 - 4b_0}}{2} \ .$$

On the other hand, by the factor theorem we have

$$x^2 + b_1x + b_0 = (x - r_1)(x - r_2)$$

whence

$$b_1 = -(r_1 + r_2)$$

$$b_0 = r_1r_2 \ .$$

Thus we have two sets of relations between the roots and the coefficients of the given monic quadratic equation, the first set consisting of expressions of the roots in terms of the coefficients and the second set consisting of expressions of the coefficients in terms of the roots.

Similarly given a cubic equation

$$x^3 + c_2x^2 + c_1x + c_0 = 0$$

we also have two such sets of relations. Now the expressions of the roots $r_1, r_2$ and $r_3$ in terms of the coefficients $c_0, c_1$ and $c_2$ constitute the substance of Cardano's formulae in 4.3.1 which are too complicated to be reproduced here. To obtain the second set of relations we use the factorization

$$x^3 + c_2x^2 + c_1x + c_0 = (x - r_1)(x - r_2)(x - r_3) \ .$$

After expanding the right-hand side and comparing corresponding terms, we get

$$\begin{aligned} c_2 &= -(r_1 + r_2 + r_3) \\ c_1 &= r_1r_2 + r_1r_3 + r_2r_3 \\ c_0 &= -r_1r_2r_3 \ . \end{aligned}$$

Given an equation $f(x) = 0$, to have the first set of relations expressing the roots in terms of the coefficients amounts to a complete solution of the proposed equation. This is, therefore, not always possible. For example, we would not have such a set of relations for a quintic equation with general coefficients. We shall show that it is always possible to get the second set of relations which, under certain circumstances, may even lead to a complete solution of the equation.

Let

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$$

be a monic equation with real coefficients. By the fundamental theorem of algebra and the factor theorem, the monic polynomial $f(x)$ has a (real or complex) root $r_1$ and can be factorized into $f(x) = (x-r_1)f_1(x)$ where $f_1(x)$ is a monic polynomial of degree $n-1$. For the same reason $f_1(x)$ has a root $r_2$ and we get $f(x) = (x-r_1)(x-r_2)f_2(x)$. Further factorization will lead finally to the complete factorization of $f(x)$:

$$f(x) = (x - r_1)(x - r_2) \cdots (x - r_n)$$

where the numbers $r_1, r_2, \ldots, r_n$ , not necessarily all distinct, are the

$n$ roots of the equation $f(x) = 0$. After expanding the right-hand side, we get

$$\begin{aligned} f(x) &= x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0 \\ &= x^n - (r_1 + \cdots + r_n)x^{n-1} + (r_1r_2 + \cdots + r_{n-1}r_n)x^{n-2} + \cdots \\ &\quad + (-1)^n r_1r_2\cdots r_n \ . \end{aligned}$$

Hence a comparison of the corresponding terms will yield:

$$\begin{aligned} -a_{n-1} &= r_1 + r_2 + \cdots + r_n \\ a_{n-2} &= r_1r_2 + r_1r_3 + \cdots + r_{n-1}r_n \\ -a_{n-3} &= r_1r_2r_3 + r_1r_2r_4 + \cdots + r_{n-2}r_{n-1}r_n \\ &\quad \cdots\cdots\cdots\cdots \\ (-1)^n a_n &= r_1r_2\cdots r_n \end{aligned}$$

which is the second set of relations.

We have therefore established the following theorem on the basic relation between the roots and the coefficients of an equation.

5.1.1 THEOREM. *In an equation in the unknown $x$ of degree $n$, in which the leading coefficient is one, the sum of $n$ roots equals the negative of the coefficient of $x^{n-1}$, the sum of the $\binom{n}{2}$ products of roots two at a time equals the coefficients of $x^{n-2}$, the sum of the $\binom{n}{3}$ products of roots three at a time equals the negative of the coefficient of $x^{n-3}$, etc.; finally the product of the $n$ roots equals the constant term or its negative according as $n$ is even or odd.*

The expressions of the coefficients of a quadratic and a cubic equation are given earlier. The coefficients of a quartic equation

$$x^4 + d_3x^3 + d_2x^2 + d_1x + d_0 = 0$$

in terms of its roots $r_1, r_2, r_3$ and $r_4$ are therefore as follows:

$$\begin{aligned} -d_3 &= r_1 + r_2 + r_3 + r_4 \\ d_2 &= r_1r_2 + r_1r_3 + r_1r_4 + r_2r_3 + r_2r_4 + r_3r_4 \\ -d_3 &= r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + r_2r_3r_4 \\ d_0 &= r_1r_2r_3r_4 \ . \end{aligned}$$

A general formula of the coefficient in terms of the roots can be given as follows

$$(-1)^k a_{n-k} = \sum r_{i_1} r_{i_2} \cdots r_{i_k}$$

where the summation is taken over all $\binom{n}{k}$ products of the roots $k$ at a time. One convenient way of writing the terms of the summation is to arrange the indices $i_j$ in a strictly increasing order, i.e. $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. For example the product $r_5 r_7 r_4 r_2 r_8$ should be written into $r_2 r_4 r_5 r_7 r_8$. Secondly we adopt a lexicographic order for the individual terms of the sum as we have done so for the coefficients of the quadratic, cubic and quartic cases.

The relations provided by Theorem 5.1.1 enables us to write down an equation in terms of the known or unknown roots. As a matter of fact this is precisely what we did at one stage in the derivation of Cardano's formulae. There we used the known relations ${z_1}^3 + {z_2}^3 = -q$ and $27{z_1}^3{z_2}^3 = -p^3$ between the unknown quantities ${z_1}^3$ and ${z_2}^3$ to write the quadratic equation

$$(z^3)^2 + qz^3 - \frac{1}{27}p^3 = 0$$

in $z^3$. By themselves the relations of Theorem 5.1.1 would not lead us to a solution of the given equation. However we shall see in the following examples that they can prove to be very useful when used in conjunction with some extra information on the roots.

5.1.2 EXAMPLE. *Solve $x^3 - 5x^2 + 8x - 4 = 0$ given that two of its roots are equal.*

SOLUTION: Let $r_1, r_2, r_3$ be the three unknown roots of the equation. The extra information is that two of them are identical, say $r_1 = r_2$. The three relations between roots and coefficients are then

$$\begin{aligned} 2r_1 + r_3 &= 5 \\ {r_1}^2 + 2r_1 r_3 &= 8 \\ {r_1}^2 r_3 &= 4 \ . \end{aligned}$$

The first two relations yield $r_1 = r_2 = 2$ and $r_3 = 1$ or $r_1 = r_2 = \frac{4}{3}$ and $r_3 = \frac{7}{3}$. The first set of values of $r_1, r_2, r_3$ are seen to satisfy the third

relation. Therefore the three roots of the given equations are $2, 2, 1$. It is not necessary to test the other set of values, since the given equation can not have two different sets of roots.

5.1.3 EXAMPLE. *Solve* $x^4 - 2x^3 - 21x^2 + 22x + 40 = 0$, *given that its roots are in arithmetic progression.*

SOLUTION: The extra information allows us to write the four roots of the given equation as $\alpha - 3\delta$, $\alpha - \delta$, $\alpha + \delta$, $\alpha + 3\delta$ with unknowns $\alpha$ and $\delta$. To find two unknowns we usually only need two relations among them. Choose, for example,

$$\begin{aligned} 2 &= (\alpha - 3\delta) + (\alpha - \delta) + (\alpha + \delta) + (\alpha + 3\delta) = 4\alpha \\ -21 &= (\alpha - 3\delta)(\alpha - \delta) + (\alpha - 3\delta)(\alpha + \delta) + (\alpha - 3\delta)(\alpha + 3\delta) \\ &\qquad + (\alpha - \delta)(\alpha + \delta) + (\alpha - \delta)(\alpha + 3\delta) + (\alpha + \delta)(\alpha + 3\delta) \\ &= 6\alpha^2 - 10\delta^2 \, . \end{aligned}$$

We find $\alpha = \frac{1}{2}$ and $\delta = \pm\frac{3}{2}$. Both values of $\delta$ yield the arithmetic progression $-4, -1, 2, 5$. To ascertain that they are the roots of the given equation, we may either verify the remaining two relations between roots and coefficients or verify the given equation by substitutions.

5.1.4 REMARKS. The observant reader may have noticed that in the examples we do not need the full set of $n$ relations of Theorem 5.1.1 to find the roots. The reason for this is that the extra information on the roots yields one or more relations between the roots and that these extra relations are being absorbed into the simultaneous equations. For example, in Example 5.1.2, $r_1 = r_2$ is taken into account when we write $2r_1 + r_3 = 5$ and ${r_1}^2 + 2r_1r_3 = 8$. In Example 5.1.3, the four unknowns $r_1, r_2, r_3, r_4$ are reduced into two unkowns $\alpha$ and $\delta$ and this reduction is absorbed when we write $2 = 4\alpha$ and $-12 = 6\alpha^2 - 10\delta^2$. On the other hand the extra information on the roots of a problem may turn out to be incompatible with the actual roots of the given equation and hence also incompatible with the $n$ relations of Theorem 5.1.1. In this case the simultaneous equations in $r_1, r_2, \ldots$ will be unsolvable. Thus the only possible conclusion is that the problem is not well-posed and has no solution.

Another kind of application of Theorem 5.1.1 can be found in the following examples.

5.1.5 EXAMPLE. *Show that if $\omega$ is an imaginary cube root of unity then $\omega^2 + \omega + 1 = 0$.*

PROOF: Being a cube root of unity, $\omega^3 = 1$ and $(\omega^2)^3 = (\omega^3)^2 = 1$. It follows from $\omega - \omega^2 = \omega(1-\omega)$ and $\omega$ being imaginary that $\omega \neq \omega^2$. Therefore $1, \omega$ and $\omega^2$ are the three roots of the cubic equation $x^3 - 1 = 0$. Hence $\omega^2 + \omega + 1 = 0$ by 5.1.1.

5.1.6 EXAMPLE. *Show that if*

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = (x-c_1)(x-c_2)\cdots(x-c_n)$$

*then*

$$\begin{aligned}&(1 + a_{n-2} + a_{n-4} + a_{n-6} + \cdots)^2 - (a_{n-1} + a_{n-3} + a_{n-5} + \cdots)^2 \\ =&(1-{c_1}^2)(1-{c_2}^2)\cdots(1-{c_n}^2)\ .\end{aligned}$$

PROOF: Let

$$\begin{aligned}f(x) &= x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0, \\ g(x) &= x^n + a_{n-2}x^{n-2} + a_{n-4}x^{n-4} + \cdots \quad \text{and} \\ h(x) &= a_{n-1}x^{n-1} + a_{n-3}x^{n-3} + \cdots\ .\end{aligned}$$

Then $f(x) = g(x) + h(x)$ . Substituting $-x$ for $x$ we get

$$\begin{aligned}(-1)^n f(-x) &= (x+c_1)(x+c_2)\cdots(x+c_n) \\ (-1)^n f(-x) &= g(x) - h(x)\ .\end{aligned}$$

Therefore we have

$$\begin{aligned}(-1)^n f(-x)f(x) &= (x-{c_1}^2)(x-{c_2}^2)\cdots(x-{c_n}^2) \\ (-1)^n f(-x)f(x) &= g(x)^2 - h(x)^2\ .\end{aligned}$$

Thus

$$g(x)^2 - h(x)^2 = (x-{c_1}^2)(x-{c_2}^2)\cdots(x-{c_n}^2)\ .$$

Substituting 1 for $x$ in the last identity, we get

$$\begin{aligned}&(1+a_{n-2}+a_{n-4}+\cdots)^2-(a_{n-1}+a_{n-3}+\cdots)^2\\=&(1-c_1{}^2)(1-c_2{}^2)\cdots(1-c_n{}^2)\ .\end{aligned}$$

5.1.7 REMARKS. Before we move to another topic, let us made a general observation on Theorem 5.1.1. If we regard the roots $r_1, r_2, \ldots, r_n$ of $f(x) = 0$ as $n$ unknowns, then the theorem readily provides us with a set of $n$ simultaneous equations in these $n$ unknowns. At first sight it might suggest that these simultaneous equations could provide us with an alternative way to solve $f(x) = 0$ . Such, however, is not the case. Let us consider, for example, the case of a cubic equation

$$x^3+ax^2+bx+c=0\ .$$

Denoting the three roots by $\alpha, \beta, \gamma$ we get

$$\begin{aligned}-a&=\alpha+\beta+\gamma\\b&=\alpha\beta+\alpha\gamma+\beta\gamma\\-c&=\alpha\beta\gamma\ .\end{aligned}$$

The usual method to solve this set of equations in $\alpha$, $\beta$ and $\gamma$ is by elimination. Thus we multiply the first by $-\alpha^2$, the second by $\alpha$, the third by $-1$ and add up to get

$$a\alpha^2+b\alpha+c=-\alpha^3$$

i.e.

$$\alpha^3+a\alpha^2+b\alpha+c=0\ .$$

But this is clearly just the original cubic equation, now written in the unknown $\alpha$ instead of $x$. Thus, in the absence of extra information on the roots, no advantage can be gained by the relations of Theorem 5.1.1 alone.

## EXERCISE 5A

1. If the sum of two of the roots of $x^3-x^2-2x+2=0$ is zero, find all the roots.

2. In Question 1, it can be noted that the constant term and the coefficients of $x$ and $x^2$ satisfy $(-1)(-2) - 2 = 0$. In general, prove that if the sum of two of the roots of $x^3 + px^2 + qx + r = 0$ is zero, then $pq - r = 0$.
3. Solve $x^4 - 10x^3 + 15x^2 + 50x - 56 = 0$, given that the roots are in arithmetic progression.
4. The roots of the equation
$$ax^3 + 3bx^2 + 3cx + d = 0$$
are in arithmetic progression. Express $c$ in terms of $a$, $b$, and $d$.
5. Solve $15x^3 - 23x^2 + 9x - 1 = 0$, given that the roots are in harmonic progression.
6. If two of the roots of $x^3 - 7x^2 + 16x - 12 = 0$ are in the ratio 3 to 2, find all the roots.
7. Solve $x^4 - 2x^3 - 11x^2 + 12x + 36 = 0$, given that there are 2 pairs of equal roots.
8. The sum of two of the roots of
$$x^3 + px^2 + p^2x + r = 0$$
is 1. Prove that $v = (p + 1)(p^2 + p + 1)$.
9. Let $\alpha$, $\beta$ and $\gamma$ be the roots of $x^3 + px^2 + qx + r = 0$. Find a cubic equation with roots $\beta\gamma$, $\gamma\alpha$ and $\alpha\beta$.
10. If $\alpha$, $\beta$ and $\gamma$ are the roots of $x^3+px^2+qx+r = 0$, find a cubic equation with roots $\dfrac{\alpha}{\beta\gamma}$, $\dfrac{\beta}{\alpha\gamma}$, $\dfrac{\gamma}{\alpha\beta}$.
11. Let $\alpha$, $\beta$ be real roots of $x^2 + px + q = 0$, where $\alpha < \beta$. If $m$ is a real number such that $\alpha < m < \beta$, show that $m^2 + pm + q < 0$. Can you give a geometric interpretation to the result?
12. If $\alpha$, $\beta$ are the roots of $ax^2 + bx + c = 0$ $(ac \neq 0)$, and $\alpha + \beta$, $\alpha\beta$ are the roots of $ax^2 - bx + c = 0$. Find $\alpha$ and $\beta$.
13. Let $\alpha$, $\beta$ be real roots of $x^2 + 2(a + 3)x + 2a + 4 = 0$ and $a$ is a real number. Prove that $(\alpha - 1)^2 + (\beta - 1)^2$ attains its minimum when $a = -3$.
14. Given that $x^3 + px + q = 0$ is a polynomial over $\mathbf{R}$ and the complex number $a+bi$ is one of its roots. Show that $2a$ is a root of $x^3+px-q = 0$.

15. Find a real value of $k$ such that $x^3 + kx^2 + 3 = 0$ has a root equal to the sum of the other two. Hence, solve the equation for this value of $k$.

16. If $\alpha$, $\beta$, and $\gamma$ are the roots of $x^3 + px^2 + r = 0$, find a cubic equation whose roots are $\alpha^2$, $\beta^2$, and $\gamma^2$, without using Theorem 5.1.1. [Hint: Use a transformation $y = x^2$.] Hence, find the values of $\alpha^2 + \beta^2 + \gamma^2$ and $\beta^2\gamma^2 + \gamma^2\alpha^2 + \alpha^2\beta^2$ in terms of $p$ and $r$.

17. If the difference of the roots of $x^2 + px + q = 0$ is equal to the difference of the roots of $x^2 + qx + p = 0$, show that $p = q$ or $p + q = -4$.

18. Given that $\alpha$ and $\beta$ are the roots of $x^2 + px + 1 = 0$ and, $\gamma$ and $\delta$ are the roots of $x^2 + qx + 1 = 0$. Prove that $(\alpha - \gamma)(\beta - \gamma)(\alpha + \delta)(\beta + \delta) = q^2 - p^2$.

19. If the lengths of the three sides of a triangle are the roots of $x^3 - px^2 + qx - r = 0$, show that the area of this triangle is $\frac{1}{4}\sqrt{4p^2q - p^4 - 8rp}$. [Hint: Do you remember any special formula for areas of triangles?]

20. If $\alpha_n$ and $\beta_n$ are the roots of $x^2 + (2n+1)x + n^2 = 0$, for $n \in \mathbf{N}$, find the value of

$$\frac{1}{(\alpha_3 + 1)(\beta_3 + 1)} + \frac{1}{(\alpha_4 + 1)(\beta_4 + 1)} + \cdots + \frac{1}{(\alpha_{20} + 1)(\beta_{20} + 1)} .$$

21. Find the values of $m$ such that the sum of the roots of the equation

$$3x^2 - (4m^2 - 1)x + m(m-2) = 0$$

is equal to the sum of the reciprocals of the roots. Hence solve the equation for these values of $m$.

22. If the real polynomial equation $x^4 - 6x^3 + ax^2 + bx + 2 = 0$ has four real roots, prove that at least one of the roots is less than 1.

23. Given that the real polynomial equation $f(x) = x^4 + px^3 + qx^2 + rx + s = 0$ has two pairs of equal roots, where $p > 0$. Show that the roots of $f(x) = 0$ are also roots of $2px^2 + p^2x + 2r = 0$. Hence, find the condition that the roots of $f(x) = 0$ are real.

24. It is known that $\tan\theta$ and $\tan(\frac{\pi}{4} - \theta)$ are the roots of the equation $x^2 + px + q = 0$ and that the roots are in the ratio 3 to 2, where $q \neq 1$.
    (a) Show that $q - p - 1 = 0$.
    (b) Find the values of $p$ and $q$.

25. Let $\alpha$, $\beta$, $\gamma$, and $\delta$ be the roots of $x^4 + px^3 + qx^2 + rx + s = 0$.
    (a) If $\alpha\beta = \gamma\delta$, show that $p^2 s - r^2 = 0$.
    (b) If $\alpha + \beta = \gamma + \delta$, show that $p^3 - 4pq + 8r = 0$.

26. Given that $\alpha$, $\beta$, $\gamma$ and $\delta$ are the roots of the equation

$$px^4 + x^3 + (p+q)x^2 - x + q = 0 ,$$

where $p \neq 0$ and $\alpha + \beta = 0$.
    (a) Show that $p + q = 0$.
    (b) Show that $\alpha$ and $\beta$ are roots of $x^2 - 1 = 0$ and hence express $\gamma$ and $\delta$ in terms of $p$.

27. Let $\alpha$, $\beta$ and $\gamma$ be the roots of the equation

$$x^3 + qx - r = 0, \quad \text{where} \quad r \neq 0 .$$

    (a) Find a cubic equation whose roots are $\alpha^3$, $\beta^3$ and $\gamma^3$.
    (b) Show that $\dfrac{\alpha}{\beta} + \dfrac{\beta}{\alpha} = \dfrac{\gamma^3}{r} - 2$.
    (c) By using (a) and (b), or otherwise, find the cubic equation with roots $\dfrac{\alpha}{\beta} + \dfrac{\beta}{\alpha}$, $\dfrac{\beta}{\gamma} + \dfrac{\gamma}{\beta}$ and $\dfrac{\gamma}{\alpha} + \dfrac{\alpha}{\gamma}$.

28. Let $\alpha$, $\beta$, $\gamma$ and $\delta$ be the roots of $x^4 + px^3 + qx^2 + rx + s = 0$ and $\alpha + \beta = \gamma + \delta$.
    (a) Show that $\alpha\beta + \gamma\delta = q - \dfrac{p^2}{4}$.
    (b) (i) By using (a), show that $\alpha\beta$ and $\gamma\delta$ are the roots of $x^2 - (q - \dfrac{p^2}{4})x + s = 0$; and hence
        (ii) find a quadratic equation that has roots $\alpha$ and $\beta$ and another quadratic equation that has roots $\gamma$ and $\delta$, the coefficients of both equations being expressed in terms of $p$, $q$ and $s$.

29. (a) If $\alpha$, $\beta$ and $\gamma$ are the roots of the equation $x^3 + px^2 + qx + r = 0$, find the values of $\alpha + \beta + \gamma$, $\alpha^2 + \beta^2 + \gamma^2$, and $\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2$ in terms of $p$, $q$, and $r$.

(b) By using (a), solve

$$\begin{cases} x + y + z = -2 \\ x^2 + y^2 + z^2 = 6 \\ x^2y^2 + y^2z^2 + z^2x^2 = 9 \,. \end{cases}$$

30. With reference to Question 29(a), find also $\alpha^3 + \beta^3 + \gamma^3$ and hence solve

$$\begin{cases} x + y + z = 5 \\ x^2 + y^2 + z^2 = 9 \\ x^3 + y^3 + z^3 = 17 \,. \end{cases}$$

31. Given that

$$\begin{cases} x + ay + a^2z + a^3 = 0 \\ x + by + b^2z + b^3 = 0 \\ x + cy + c^2z + c^3 = 0 \,, \end{cases}$$

express $x$, $y$, and $z$ in terms of $a$, $b$ and $c$.

32. Given that $r_1, r_2, \ldots, r_n$ are the roots of the equation

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \,,$$

find

(a) $\sum_{i=1}^{n} r_i^2$

(b) $\sum_{i=1}^{n} \dfrac{1}{r_i}$

(c) $\sum \dfrac{1}{r_{i_1} r_{i_2}}$

(d) $\sum \dfrac{1}{r_{i_1} r_{i_2} \cdots r_{i_k}}$

(e) $\sum \dfrac{r_{i_1}^2 + r_{i_2}^2}{r_{i_1} r_{i_2}}$

(f) $\sum (r_{i_1} - r_{i_2})^2 r_{i_3} r_{r_4} \cdots r_{i_n}$.

33. Let $\alpha$, $\beta$ and $\gamma$ be the roots of the equation

$$x^3 + 3px + q = 0, \quad \text{where } q \neq 0 \,.$$

(a) Express $\alpha^2 + \beta^2 + \gamma^2$ in terms of $p$.

(b) Show that $(\beta-\gamma)^2 = -6p - \alpha^2 + \frac{2q}{\alpha}$ and $\alpha = \frac{3q}{(\beta-\alpha)^2+3p}$.

(c) Show that $(\beta-\gamma)^2$ is a root of the equation

$$y^3 + 18py^2 + 81p^2y + 27(q^2+4p^3) = 0 .$$

Hence, show that the condition for the equation $x^3 + 3px + q = 0$ to have two equal roots is $q^2 + 4p^3 = 0$.

34. As a generalization to Question 24, we change the condition that the roots are in the ratio $m$ to $n$, where $m = (2h-1)(4h-1)$ and $n = 2h$, $h$ is any positive integer (for $h = 1$, we are back to Question 24).

(a) (i) Show that $mnt^2 + (m+n)t - 1 = 0$.

(ii) Express the discriminant of the equation in (i) in terms of $h$.

(b) By (a), solve the equation $x^2 + px + q = 0$.

35. Given that the roots $\alpha$, $\beta$ and $\gamma$ of $x^3 + px^2 + qx + r = 0$ are all real.

(a) If $\alpha$, $\beta$ and $\gamma$ are the lengths of the sides of a triangle, then show that

$$p < 0, \quad q > 0, \quad r < 0, \quad \text{and}$$
$$p^3 > 4pq - 8r .$$

(b) (i) Suppose $p < 0$, $q > 0$ and $r < 0$, show that $\alpha$, $\beta$ and $\gamma$ are all positive.

(ii) Hence, with the conditions in (i) and $p^3 > 4pq - 8r$, show that $\alpha$, $\beta$, and $\gamma$ are the lengths of the sides of a triangle.

## 5.2 Integral roots

In the remaing sections of this chapter we shall investigate some other relations between roots and coefficients, and study appropriate methods for solving certain types of equation.

The last of the $n$ relations between the roots and the coefficients

$$r_1r_2\cdots r_n = \pm a_0$$

given in Theorem 5.1.1 would suggest that the roots of a monic equa-

tion are divisors of the constant term. This is certainly true if we know beforehand that all the roots and the constant terms are integers. In this case we simply proceed to single out the roots among the divisors of $a_0$. Unfortunately, it is seldom that we could have such information on the roots before the equation is actually solved. The same conclusion may be quite wrong if not all the roots are integers. Take, for example, the equation

$$x^2 - \frac{9}{2}x + 2 = 0 .$$

The roots are 4 and $\frac{1}{2}$, and it would be wrong to say that 2 is divisible by 4 just because $2 = 4 \times \frac{1}{2}$.

Instead of working with the hypothesis that the roots are integers, we consider the case in which the coefficients are all integers. Let us first prove a simple but very useful theorem for equations with integral coefficients.

5.2.1 THEOREM. *For an equation whose coefficients are all integers any integral root is a divisor of the constant term.*

PROOF: The theorem may be restated as follows. *If $r \in \mathbf{Z}$ is a root of the equation*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

*where $a_i \in \mathbf{Z}$ for $i = 0, 1, \ldots, n$, then $r | a_0$.* The proof is simple and does not depend on 5.1.1. By hypothesis $a_n r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0 = 0$. Therefore $a_0 = -r(a_n r^{n-1} + a_{n-1}r^{n-2} + \cdots + a_2 r + a_1)$. Since $a_0$ and both factors on the right-hand side are integers, $a_0$ is divisible by $r$.

Given an equation with integral coefficients, it is natural that we should find out its integral roots first and other types of roots later. The theorem now gives definite indication as to where the integral roots could be found. Suppose that we have made a good guess that a divisor $r$ of $a_0$ could be an integral root. Then we would certainly proceed to verify it by a synthetic substitution:

| | | | | | |
|---|---|---|---|---|---|
| $a_n$ | $a_{n-1}$ | $a_{n-2}$ | $\cdots$ | $a_1$ | $a_0$ |
| | $rd_n$ | $rd_{n-1}$ | $\cdots$ | $rd_2$ | $rd_1$ |
| $d_n = a_n$ | $d_{n-1}$ | $d_{n-2}$ | $\cdots$ | $d_1$ | $0$ |

Thus without doing any extra work we also obtain the factorization $f(x) = (x - c)(d_n x^{n-1} + d_{n-1}x^{n-2} + \cdots + d_2 x + d_1)$ where $d_i \in \mathbf{Z}$ . Therefore Theorem 5.2.1 can be applied to the equation

$$d_n x^{n-1} + d_{n-1}x^{n-2} + \cdots + d_2 x + d_1 = 0$$

to find further integral roots of the given equation $f(x) = 0$ .

5.2.2 EXAMPLE. *Solve the equation*

$$x^4 + 2x^3 - 21x^2 - 22x + 40 = 0$$

SOLUTION: Since we are primarily interested in the integral roots of the equation, the most obvious thing to do on the outset is to test whether 0, 1 or $-1$ is a root. By an inspection of the coefficients we see that 1 is a root. A synthetic substitution produces

| | | | | |
|---|---|---|---|---|
| 1 | 2 | $-21$ | $-22$ | 40 |
| | 1 | 3 | $-18$ | $-40$ |
| 1 | 3 | $-18$ | $-40$ | 0 |

and $f(x) = (x - 1)(x^3 + 3x^2 - 18x - 40)$ . Further roots of $f(x) = 0$ must be found among those of the equation

$$x^3 + 3x^2 - 18x - 40 = 0 \ .$$

Now none of the values 0, 1 and $-1$ is a root. By Theorem 5.2.1 the integral roots, if any, are among the divisors $\pm 2$, $\pm 4$, $\pm 5$, $\pm 8$, $\pm 10$, $\pm 20$, $\pm 40$ of the constant terms $-40$. We can exclude $\pm 10$, $\pm 20$, $\pm 40$ since the leading term $x^3$ at such values will outweigh all other terms. We try $-2$ and obtain

| | | | |
|---|---|---|---|
| 1 | 3 | $-18$ | $-40$ |
| | $-2$ | $-2$ | 40 |
| 1 | 1 | $-20$ | 0 |

Thus $-2$ is a root and the remaining roots are found by solving $x^2+x-20=0$. Now $x^2+x-20=(x-4)(x+5)$. Therefore the four roots of the equation are $-5,-2,1,4$.

5.2.3 EXAMPLE. *Solve the equation*

$$x^4+6x^3-18x^2-19x-24=0\ .$$

SOLUTION: For $r=3$ we have

$$\begin{array}{rrrrr} 1 & 6 & -18 & -19 & -24 \\ & 3 & 27 & 27 & 24 \\ \hline 1 & 9 & 9 & 8 & 0 \end{array}$$

For $s=-8$ we have

$$\begin{array}{rrrr} 1 & 9 & 9 & 8 \\ & -8 & -8 & -8 \\ \hline 1 & 1 & 1 & 0 \end{array}$$

Now $x^2+x+1=0$ has no real roots. By Example 5.1.5 the roots of this equation are the imaginary cube roots of unity. Thus $-8, 3, -\frac{1}{2}(1+i\sqrt{3})$ and $-\frac{1}{2}(1-i\sqrt{3})$ are the four roots of the given equation.

The amount of calculation in an application of Theorem 5.2.1 depends on the number of divisors of the constant term. If the constant term $a_0$ is 1 or $-1$ then it has only two divisors 1 and $-1$ and the application of Theorem 5.2.1 would be very simple. If $a_0$ is a prime number $p$, then we would carry out a test by synthetic substitution on the four divisors $\pm 1$ and $\pm p$ which is still easy to handle. However if $a_0$ is a composite number with $s$ different prime factors, then $a_0$ will have no less than $2^{s+1}$ divisors. The amount of work to be done in the test could be quite considerable. Therefore it would be very desirable to have some means to reduce the number of divisors to be tested. One such is provided by the following theorem.

5.2.4 THEOREM. *Let* $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$ *be an equation with integral coefficients and let* $b$ *be an integer. Then* $b$ *is not*

*a root of $f(x) = 0$ if an integer $m \neq 0$ can be found such that $f(m) \neq 0$ is not divisible by $b - m$ .*

PROOF: Let $m$ be an integer such that $f(m)$ is not divisible by $b - m$. Suppose to the contrary that $b$ is a root of the equation $f(x) = 0$ . Then $f(x) = (x-b)q(x)$, where $q(x) = d_n x^{n-1} + d_{n-1}x^{n-1} + \cdots + d_2 x + d_1$ also has integral coefficients. Therefore it would follow that $f(m) = (m - b)q(m)$, where all three expressions $f(m)$, $m - b$ and $q(m)$ are integers. But this would mean that $f(m)$ is divisible by $m - b$ which is impossible. We must therefore conclude that $b$ is not a root of the given equation $f(x) = 0$.

To make use of Theorem 5.2.4 we may choose any integer $m \neq 0$ as long as $f(m) \neq 0$. In general we would prefer a small non-zero value of $|f(m)|$ so that many divisors $b$ of the constant term $a_0$ could be quickly eliminated as possible roots of the equation.

5.2.5 EXAMPLE. *Solve the equation*

$$x^4 + 9x^3 + 24x^2 + 23x + 15 = 0 \ .$$

SOLUTION: Since all coefficients are positive the equation cannot have positive roots. The divisors of 15 to be tested are $-1, -3, -5, -15$. The factor $-1$ can be eliminated as a root since the alternating sum of the coefficients is non-zero. In fact $f(-1) = 8$ (i.e. $f(m) = 8$ with $m = -1$) which is not divisible by $-15 + 1 = -14$ (i.e. $b - m = -14$ with $b = -15$). Therefore $-15$ can be eliminated as a root of the equation. For the divisor $-3$, we get

$$\begin{array}{ccccc} 1 & 9 & 24 & 23 & 15 \\ & -3 & -18 & -18 & -15 \\ \hline 1 & 6 & 6 & 5 & 0 \end{array}$$

For the divisor $-5$, we get

$$\begin{array}{cccc} 1 & 6 & 6 & 5 \\ & -5 & -5 & -5 \\ \hline 1 & 1 & 1 & 0 \end{array}$$

Therefore $f(x) = (x+3)(x+5)(x^2 + x + 1)$ . The roots of the equation are $-5, -3, \omega$ and $\omega^2$.

5.2.6 EXAMPLE. *Solve the equation*

$$f(x) = x^3 - 2x^2 - 23x + 60 = 0 .$$

SOLUTION: $f(1) = 36$ . Among the divisors of 60, we can exclude $-4$, $\pm 6$, $-10$, $\pm 12$, $\pm 15$, $\pm 20$, $\pm 30$, $\pm 60$. Further tests lead to $f(x) = (x-4)(x-3)(x+5)$. Therefore the roots of the equation are $-5, 3, 4$.

5.2.7 REMARKS. In the last example we have a relatively large constant term 60. In such case we advise our readers to proceed in a more orderly manner so that no divisor of 60 is omitted. To begin, we write $60 = 2\times 2\times 3\times 5$ as a product of primes. Then all divisors $b$ of 60 are just partial products of these prime factors together with their negatives. For the elimination test according to Theorem 5.2.4, we use the values $m = 1$ and $f(m) = 36$. Thus we set up a table for all the possible values of $b$ and $b - m$ as follows:

$$m = 1; \quad f(m) = 36$$

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $b-m$ | 0 | 1 | 2 | 3 | 4 | 5 | 9 | 11 | 14 | 19 | 29 | 39 |
| $b$ | 1 | 2 | 3 | 4 | 5 | 6 | 10 | 12 | 15 | 20 | 30 | 60 |
| $b$ | −1 | −2 | −3 | −4 | −5 | −6 | −10 | −12 | −15 | −20 | −30 | −60 |
| $b-m$ | −2 | −3 | −4 | −5 | −6 | −7 | −11 | −13 | −16 | −21 | −31 | −61 |

Finally we just eliminate all the numbers on the top and the bottom rows that are not divisors of $f(m) = 36$ and carry out a test on the remaining ones.

## EXERCISE 5B

1. Solve the following equations.
   (a) $x^3 + 2x^2 - x - 2 = 0$.
   (b) $x^3 - 3x^2 - 9x - 5 = 0$.
   (c) $x^3 + 2x^2 - 5x - 6 = 0$.
   (d) $2x^3 - 21x^2 + 49x + 30 = 0$.
   (e) $5x^3 - 81x^2 + 316x - 60 = 0$.
   (f) $x^5 - 7x^4 + 12x^3 - 8x^2 + 56x - 96 = 0$.

2. Let $m$ be an integer and that the roots of $2x^4 + mx^2 + 8 = 0$ are all integers. Solve the equation and find the value of $m$.
3. If $a$ and $b$ are integers, prove that both of the equations

$$x^2 + 10ax + 5b + 3 = 0, \quad \text{and} \quad x^2 + 10ax + 5b - 3 = 0\ ,$$

cannot have integral roots.

4. Let $x^3 - kx^2 + kx + 15 = 0 \in \mathbf{Z}[x]$. If all the three roots of the equation are integers with two of them positive, and the sums of every two roots are in geometric progression, find all the roots and the value of $k$.
5. If $d \neq \pm 1$ is an integral root of $f(x) \in \mathbf{Z}[x]$, show that $\frac{f(1)}{\alpha-1}$ and $\frac{f(-1)}{\alpha+1}$ are integers.
6. Let $f(x) \in \mathbf{Z}[x]$. Prove that if $f(0)$ and $f(1)$ are both odd integers, then $f(x)$ cannot have integral root.
7. Let positive integers $\alpha$, $\beta$ and $\gamma$ be roots of $x^3 - 11x^2 + mx - 36 = 0$. If $\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 1$, find the value of $m$ and hence solve the given equation.
8. As a generalization to Question 6, show that for $f(x) \in \mathbf{Z}[x]$, if there are an even integer $a$ and an odd integer $b$ such that $f(a)$ and $f(b)$ are both odd, prove that $f(x)$ has no integral root.
9. If $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, where $n$ is an even integer and $a_0, a_1, \ldots, a_{n-1}$ are odd integers, show that $f(x)$ has no integral root.
10. For $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ of $\mathbf{Z}[x]$ with $a_0 \neq 0$, prove that a necessary and sufficient condition for $f(x)$ to have integral root is that there are $2(n-1)$ integers $b_i$, $c_i$ $(i = 1, \ldots, n-1)$ such that

    (i) $a_i = b_i + c_i$, $i = 1, 2, \ldots, n-1$; and

    (ii) $\dfrac{1}{c_{n-1}} = \dfrac{b_{n-1}}{c_{n-2}} = \dfrac{b_{n-2}}{c_{n-3}} = \cdots = \dfrac{b_1}{a_0}$.

## 5.3 Rational roots

In this section we continue to consider equations with integral coefficients but direct our attention to their rational roots. In Theorem 5.2.1 we find a useful relation between the integral roots and

the constant terms. Parallel to this we have in the next theorem an equally useful relation between the rational roots on the one side and the leading coefficient together with the constant term on the other side.

5.3.1 THEOREM. *Let*

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$$

*be an equation with integral coefficients. If a rational number* $\frac{c}{d}$ *written in the lowest term is a root of the equation then* $c|a_0$ *and* $d|a_n$ .

PROOF: It follows from

$$a_n(\frac{c}{d})^n + a_{n-1}(\frac{c}{d})^{n-1} + \cdots + a_1(\frac{c}{d}) + a_0 = 0$$

that

$$a_nc^n = -d(a_{n-1}c^{n-1} + a_{n-2}c^{n-2}d + \cdots + a_1cd^{n-2} + a_0d^{n-1}) .$$

Since $\gcd(c,d) = 1$, we have $d|a_n$ . On the other hand

$$a_0d^n = -c(a_nc^{n-1} + a_{n-1}c^{n-2}d + \cdots + a_1d^{n-1}) .$$

Therefore $c|a_0$. The proof is complete.

5.3.2 COROLLARY. *Let*

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$$

*be an equation with integral coefficients in which the leading coefficient equals 1. Then every rational root of the equation is an integer.*

The criterion of Theorem 5.3.1 can be applied in the most straightforward manner.

5.3.3 EXAMPLE. *Find all rational roots of*

$$4x^3 - 5x^2 - 5x - 9 = 0 .$$

SOLUTION: The positive divisors of the leading coefficient are $1, 2, 4$ and those of the constant terms are $1, 3, 9$ . By Theorem 5.3.1 there are 18 possible values for the rational roots of the equation:

$$\pm 1, \pm 3, \pm 9; \ \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{9}{2}; \ \pm\frac{1}{4}, \pm\frac{3}{4}, \pm\frac{9}{4} .$$

It is easy to see that none of the integral values is a root of the equation. Try $\frac{9}{4}$

$$\begin{array}{cccc} 4 & -5 & -5 & -9 \\ & 9 & 9 & 9 \\ \hline 4 & 4 & 4 & 0 \end{array}$$

Therefore $\frac{9}{4}$ is a root and $4x^3 - 5x - 9 = 4(x - \frac{9}{4})(x^2 + x + 1)$. Therefore there is no more rational root.

5.3.4 EXAMPLE. *Find all rational roots of the equation*

$$4x^3 + 16x^2 + 21x + 9 = 0 .$$

SOLUTION: We have the same 18 possible values as in the last example. Since the coefficient of the equation are all positive, we may discard all nine positive values. The alternating sum of the coefficients is zero. Therefore $-1$ is a root. Substituting $-1$ we get

$$\begin{array}{cccc} 4 & 16 & 21 & 9 \\ & -4 & -12 & -9 \\ \hline 4 & 12 & 9 & 0 \end{array}$$

This gives an equation $4x^2 + 12x + 9 = (2x + 3)^2 = 0$. Therefore the roots of the equation are all rational and they are $-1$ and $-\frac{3}{2}$.

At first sight Corollary 5.3.2, being a special case of Theorem 5.3.1, does not offer much advantage. In the following example we shall see that it can be used in conjunction with an appropriate transformation to treat equations with rational coefficients.

5.3.5 EXAMPLE. *Find all rational roots of*

$$y^4 - \frac{40}{3}y^3 + \frac{130}{3}y^2 - 40y + 9 = 0 .$$

SOLUTION: Multiply the equation by an appropriate power of the LCM of the denominators of the coefficients to get

$$3^4y^4 - 3^3(40y^3) + 3^3(130y^2) - 3^4(40y) + 3^4(9) = 0 .$$

Replace $3y$ by $x$ to get

$$x^4 - 40x^3 + 390x^2 - 1080x + 729 = 0 .$$

Now when divided by 3 every root of the last equation in $x$ is a root of the given equation in $y$. By Corollary 5.3.2, we find that the rational roots of equation in $x$ are all integers. By an inspection of coefficients, we find that 1 is a root. Upon substitution

$$\begin{array}{rrrrr} 1 & -40 & 390 & -1080 & 729 \\ & 1 & -39 & 351 & -729 \\ \hline 1 & -39 & 351 & -729 & 0 \end{array}$$

we obtain a cubic quotient and hence we proceed to solve the equation $x^3 - 39x^2 + 351x - 729 = 0$ . The integral roots are divisors of $729 = 3^6$. Try 3

$$\begin{array}{rrrr} 1 & -39 & 351 & -729 \\ & 3 & -108 & 729 \\ \hline 1 & -36 & 243 & 0 \end{array}$$

Try 9

$$\begin{array}{rrr} 1 & -36 & 243 \\ & 9 & 243 \\ \hline 1 & -27 & 0 \end{array}$$

Therefore the roots of the equation in $x$ are $1, 3, 9, 27$. Therefore the roots of the given equation in $y$ are $\frac{1}{3}, 1, 3, 9$ .

To conclude this section we use Corollary 5.3.2 to obtain a proof of the irrationality of $\sqrt{2}$ .

5.3.6. EXAMPLE. *Prove that $\sqrt{2}$ is not a rational number.*

PROOF: By definition $\sqrt{2}$ is a root of the equation $x^2 - 2 = 0$ which has integral coefficinets. By Corollary 5.3.2, if the equation has a rational root, then it must be an integer and a divisor of 2 . In other words it must be either $\pm 1$ or $\pm 2$. But $(\pm 1)^2 - 2 \neq 0$ and $(\pm 2)^2 - 2 \neq 0$. Therefore the equation has no rational root. Now this means that as a root of the equation, $\sqrt{2}$ cannot be a rational number. The proof is complete.

## EXERCISE 5C

1. Show that the following equations have no rational root.
   (a) $3x^3 + 2x - 1 = 0$.
   (b) $2x^4 + 8x^3 + 3x^2 + 4x + 1 = 0$.
2. Find the rational roots of the following equations.
   (a) $6x^3 + 11x^2 + 6x + 1 = 0$.
   (b) $24x^5 + 10x^4 - x^3 - 19x^2 - 5x + 6 = 0$.
3. If $p$, $q$, $m$ are rational numbers such that $p = m + \frac{q}{m}$, prove that the equation $x^2 + px + q = 0$ has rational roots.
4. Given that $a$ is a positive integer and $a \neq b^2$ for any integer $b$. Show that $\sqrt{a}$ is irrational.
5. Find a quadratic equation with integral coefficients such that $2 + \sqrt{3}$ is a root. Hence, show that $2 + \sqrt{3}$ is irrational.
6. As a generalization to Question 4, suppose $a$, $b$ are integers such that $b > 0$ and $\sqrt{b}$ is not an integer. Find a quadratic equation with integral coefficients such that $a + \sqrt{b}$ is a root. Hence, show that $a + \sqrt{b}$ is irrational.
7. Prove that $\sqrt{7} - \sqrt{2}$ is irrational by using the technique as in Question 6.
8. If $p_1, \ldots, p_k$ are $k$ $(\geq 1)$ distinct positive prime integers and $n$ is any integer exceeding 1, prove that $f(x) = x^n - p_1 p_2 \ldots , p_k = 0$ has no rational root. (Thus $\sqrt[3]{6}$, $\sqrt[11]{15}$, can be classified as irrational numbers by this result immediately.)

9. (a) Show that $\tan 5\theta = \dfrac{5\tan\theta - 10\tan^3\theta + \tan^5\theta}{1 - 10\tan^2\theta + 5\tan^4\theta}$, for $\theta$ such that $\cos 5\theta \neq 0$.

   (b) Hence, or otherwise, show that $\tan\frac{\pi}{5}$ is irrational.

10. Generalize the result in Question 9 for $\tan\frac{\pi}{2n+1}$ where $n$ is a positive integer.

11. (a) Given that $a \neq \pm 2$, show that $x^{2n} - ax + 1 = 0$ has no rational root for any positive integer $n$.

    (b) Hence, show that $3x^{2n+2} - 10x^{2n+1} + 3x^{2n} - 3ax^3 + (10a+3)x^2 - (3a+10)x + 3 = 0$ has only two rational roots.

12. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbf{Z}[x]$. If a rational number $\frac{c}{d}$ written in lowest term is a root of the equation, then for any integer $k$, $(dk - c)|f(k)$.

13. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbf{Z}[x]$. If $a_0$ and $a_n$ are odd, and at least one of $f(1)$ and $f(-1)$ is odd, prove that $f(x)$ has no rational root.

14. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbf{Z}[x]$. Prove that if $a_0$, $a_n$, $f(1)$ and $f(-1)$ are not divisible by 3, then $f(x)$ has no rational root.

## 5.4 Reciprocal equations

We shall now drop the previous restriction on the coefficients and consider equations with arbitrary real coefficients again in this section. In particular we are interested in a type of equation in which the coefficients show some pattern of symmetry such as in the following equations

$$2x^5 + 3x^4 - 8x^3 - 8x^2 + 3x + 2 = 0$$
$$5x^6 - 3x^5 + 7x^4 + 0x^3 - 7x^2 + 3x - 5 = 0\ .$$

In other words, we shall study equations

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \quad (a_n \neq 0)$$

whose coefficients are *symmetric*, i.e.

$$a_n = a_0, \quad a_{n-1} = a_1, \quad a_{n-2} = a_2, \cdots$$

or whose coefficients are *skew-symmetric*, i.e.

$$a_n = -a_0, \quad a_{n-1} = -a_1, \quad a_{n-2} = -a_2, \cdots .$$

We naturally expect that some pattern of the roots will emerge from such nice pattern of the coefficients. Let us consider first the symmetric case. By the symmetry of the coefficients, we can write the polynomial $f(x)$ as

$$f(x) = a_n(x^n + 1) + a_{n-1}(x^{n-1} + x) + a_{n-2}(x^{n-2} + x^2) + \cdots$$

from which we get

$$f(x) = x^n f(\frac{1}{x}) .$$

Now it follows from $a_n \neq 0$ that $a_0 \neq 0$. Therefore 0 is not a root of $f(x) = 0$; thus every root $r$ of $f(x) = 0$ has a reciprocal $1/r$. But then

$$f(\frac{1}{r}) = \frac{1}{r^n} f(r) = 0 .$$

Therefore the reciprocal of every root of $f(x) = 0$ is a root of $f(x) = 0$.

For the skew-symmetric case we have

$$f(x) = a_n(x^n - 1) + a_{n-1}(x^{n-1} - x) + a_{n-2}(x^{n-2} - x^2) + \cdots$$

and hence

$$-f(x) = x^n f(\frac{1}{x}) .$$

Therefore the same conclusion holds.

This pattern of the roots leads us to identify a special type of equation: an equation is called a *reciprocal equation* if the reciprocal of each root is again a root. We have so far proved that if the coefficients of an equation $f(x) = 0$ are either symmetric or skew-symmetric, then $f(x) = 0$ is a reciprocal equation.

We now prove that the converse of the last statement is also true. Let us first look at some examples. The equations

$$(x - i)(x - \frac{1}{i}) = 0$$

$$3(x - 1)(x - 3)(x - \frac{1}{3}) = 0$$

are reciprocal equations. On expansion, they become

$$x^2+1=0$$
$$3x^3-13x^2+13x-3=0$$

showing a symmetric and a skew-symmetric pattern of their coefficients.

In general, let $f(x)=0$ be a reciprocal equation. We proceed to find the pattern of its cofficients. First of all, we may divide $f(x)$ by its leading coefficient and assume that

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

has a leading coefficient equal to 1. Since 0 has no reciprocal, it cannot be a root of the reciprocal equation $f(x)=0$. Therefore, $a_0\neq 0$. Suppose that $r_1,r_2,\ldots,r_n$ are the $n$ roots of $f(x)=0$. Then

$$a_0+a_1x+\cdots+a_{n-1}x^{n-1}+x^n=(x-r_1)(x-r_2)\cdots(x-r_n)\ .$$

Substituting $1/x$ for $x$ in the above equation, we get

$$\begin{aligned}
&a_0+a_1\frac{1}{x}+\cdots+a_{n-1}\frac{1}{x^{n-1}}+\frac{1}{x^n}=(\frac{1}{x}-r_1)(\frac{1}{x}-r_2)\cdots(\frac{1}{x}-r_n)\\
&=\frac{(-1)^n r_1r_2\ldots r_n}{x^n}(x-\frac{1}{r_1})(x-\frac{1}{r_2})\cdots(x-\frac{1}{r_n})\\
&=\frac{a_0}{x^n}(x-\frac{1}{r_1})(x-\frac{1}{r_2})\cdots(x-\frac{1}{r_n})\ .
\end{aligned}$$

Therefore

$$x^n+\frac{a_1}{a_0}x^{n-1}+\cdots+\frac{a_{n-1}}{a_0}x+\frac{1}{a_0}=(x-\frac{1}{r_1})(x-\frac{1}{r_2})\cdots(x-\frac{1}{r_n})\ .$$

We now put

$$g(x)=x^n+\frac{a_1}{a_0}x^{n-1}+\cdots+\frac{a_{n-1}}{a_0}x+\frac{1}{a_0}=\frac{x^n}{a_0}f(\frac{1}{x})\ .$$

Then

$$f(x)=(x-r_1)(x-r_2)\cdots(x-r_n)$$
$$g(x)=(x-\frac{1}{r_1})(x-\frac{1}{r_2})\cdots(x-\frac{1}{r_n})\ .$$

Now by the assumption that $f(x) = 0$ is a reciprocal equation, the numbers $1/r_1, 1/r_2, \dots, 1/r_n$ are just a permutation of the numbers $r_1, r_2, \dots, r_n$. Therefore

$$f(x) = g(x) = \frac{x^n}{a_0} f(\frac{1}{x}) \ .$$

A comparison of the coefficients yields

$$a_0 = \frac{1}{a_0}, \quad a_1 = \frac{a_{n-1}}{a_0}, \quad a_2 = \frac{a_{n-2}}{a_0}, \dots \ .$$

Thus in particular $a_0 = \pm 1$. We now treat the two cases separately.

If $a_0 = 1$, then $f(x) = x^n f(\frac{1}{x})$ and $a_i = a_{n-i}$; thus

$$\begin{aligned} f(x) &= x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0 \\ &= (x^n + 1) + a_1(x^{n-1} + x) + a_2(x^{n-2} + x^2) + \cdots \ . \end{aligned}$$

If $a_0 = -1$, then $-f(x) = x^n f(\frac{1}{x})$ and $a_i = -a_{n-i}$; thus

$$\begin{aligned} f(x) &= x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0 \\ &= (x^n - 1) + a_{n-1}(x^{n-1} - x) + a_{n-2}(x^{n-2} - x^2) + \cdots \ . \end{aligned}$$

In other words, the coefficients of the given reciprocal equation $f(x) = 0$ are either symmetric or skew-symmetric. We have thus proved the following theorem.

5.4.1 THEOREM. *For a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, the equation $f(x) = 0$ is a reciprocal equation if and only if*

$$f(x) = a_n(x^n \pm 1) + a_{n-1}(x^{n-1} \pm x) + a_{n-2}(x^{n-2} \pm x^2) + \cdots$$

*where either the upper signs or the lower signs hold throughout.*

Let us now study special methods for solving reciprocal equations.

5.4.2 EXAMPLE. *Solve $f(x) = x^5 + 5x^4 + 9x^3 + 9x^2 + 5x + 1 = 0$.*

SOLUTION: The given equation is a reciprocal equation of odd degree with symmetric coefficients and $f(x) = (x^5+1)+5(x^4+x)+9(x^3+x^2)$. Therefore

$-1$ is a root since $(-1)^{5-i} + (-1)^i = 0$. Division of $f(x)$ by $(x+1)$ yields the quotient

$$q(x) = x^4 + 4x^3 + 5x^2 + 4x + 1 \ .$$

We see that $q(x) = 0$ is of even degree with symmetric coefficients. Therefore it is a reciprocal equation of the form

$$(x^4 + 1) + 4(x^3 + x) + 5x^2 = 0 \ .$$

We divide the equation by $x^2$ to get

$$(x^2 + \frac{1}{x^2}) + 4(x + \frac{1}{x}) + 5 = 0 \ .$$

Next we transform this into an equation of degree 2 by putting

$$x + \frac{1}{x} = y \quad \text{thus} \quad x^2 + (\frac{1}{x})^2 = y^2 - 2$$

to obtain

$$y^2 + 4y + 3 = 0$$

which has roots $-1$ and $-3$. Substituting these values for $y$ in $x + \frac{1}{x} = y$, we get two equations

$$x + \frac{1}{x} = -1 \qquad \text{and} \qquad x + \frac{1}{x} = -3$$

in $x$. The former is $x^2 + x + 1 = 0$ which has roots $\omega$ and $\omega^2$. The latter is $x^2 + 3x + 1 = 0$ which has roots $\frac{1}{2}(-3 + \sqrt{5})$ and $\frac{1}{2}(-3 - \sqrt{5})$. Therefore the roots of the given quintic equation are $-1$, $\frac{1}{2}(-1 \pm i\sqrt{3})$, $\frac{1}{2}(-3 \pm \sqrt{5})$, one integral root, two complex roots which are conjugates and reciprocals of each other and two irrational roots which are reciprocals of each other. We remark that in the above solution, the division by $x^2$ and the multiplication by $x$ are legitimate because we are dealing with roots $x$ of a reciprocal equation which are not zero.

5.4.3 EXAMPLE. *Solve the equation* $x^6 - x^5 + x - 1 = 0$ .

SOLUTION: This is a reciprocal equation of even degree with skew-symmetric coefficients. Therefore both 1 and $-1$ are roots since $(1)^{6-i} - (1)^i = 0$ and $(-1)^{6-i} - (-1)^i = 0$. After dividing by $x^2 - 1$, we work with the

equation of the quotient

$$x^4 - x^3 + x^2 - x + 1 = 0$$

which is a reciprocal equation of even degree with symmetric coefficients and can be written as

$$(x^4 + 1) - (x^3 + x) + x^2 = 0 .$$

After division by $x^2$ we have

$$(x^2 + \frac{1}{x^2}) - (x + \frac{1}{x}) + 1 = 0 .$$

Replacing $x + \frac{1}{x}$ by $y$ as in the last example, we work with

$$y^2 - y - 1 = 0 .$$

This equation has roots $\frac{1}{2}(1 + \sqrt{5})$ and $\frac{1}{2}(1 - \sqrt{5})$. Solving the equations

$$x + \frac{1}{x} = \frac{1}{2}(1 + \sqrt{5}) \qquad \text{and} \qquad x + \frac{1}{x} = \frac{1}{2}(1 - \sqrt{5})$$

we get the roots $\frac{1}{4}(1 + \sqrt{5} \pm i\sqrt{10 - 2\sqrt{5}})$ and $\frac{1}{4}(1 - \sqrt{5} \pm i\sqrt{10 + 2\sqrt{5}})$. The roots of the given equation are therefore

$$\pm 1, \ \frac{1}{4}(1 + \sqrt{5} \pm i\sqrt{10 - 2\sqrt{5}}\,), \quad \frac{1}{4}(1 - \sqrt{5} \pm i\sqrt{10 + 2\sqrt{5}}\,) .$$

One common feature of the above examples seems to be that at some stage in the course of solution we arrive at a reciprocal equation of even degree with symmetric coefficients:

$$(x^{2m} + 1) + b_1(x^{2m-1} + x) + \cdots + b_m x^m = 0 .$$

Then after division by $x^m$ and replacement of $x + \frac{1}{x}$ by $y$ we obtain an equation of degree $m$ in the new unknown $y$. We shall see here that this is always possible.

In general given a reciprocal equation

$$f(x) = (x^n \pm 1) + a_1(x^{n-1} \pm x) + a_2(x^{n-2} \pm x^2) + \cdots = 0$$

we may distinguish four different cases:

(1) $n$ is odd $(n = 2m+1)$ and the upper signs hold.

(2) $n$ is odd $(n = 2m+1)$ and the lower signs hold.

(3) $n$ is even $(n = 2m)$ and the upper signs hold.

(4) $n$ is even $(n = 2m)$ and the lower signs hold.

Let us proceed to treat these four cases separately.

(1) In this case $f(x) = x^n f(\frac{1}{x})$ and

$$f(x) = (x^n + 1) + a_1(x^{n-1} + x) + a_2(x^{n-2} + x^2) + \cdots$$

where $n$ is odd. Therefore $-1$ is a root of $f(x) = 0$ and $x+1$ is a factor of $f(x)$. Consider the quotient

$$q(x) = \frac{f(x)}{x+1}$$

which is a polynomial of degree $n-1$. Then

$$x^{n-1}q(\frac{1}{x}) = \frac{x^{n-1}f(\frac{1}{x})}{\frac{1}{x}+1} = \frac{x^n f(\frac{1}{x})}{x+1} = \frac{f(x)}{x+1} = q(x) \ .$$

This means that $q(x)$ is a reciprocal equation in which the upper signs in 5.4.1 hold:

$$q(x) = (x^{2m} + 1) + b_1(x^{2m-1} + x) + \cdots + b_m x^n \ .$$

(2) In this case $-f(x) = x^n f(\frac{1}{x})$ and

$$f(x) = (x^n - 1) + a_1(x^{n-1} - x) + a_2(x^{n-2} - x^2) + \cdots$$

where $n$ is odd. Therefore 1 is a root of $f(x) = 0$ and $x - 1$ is a factor of $f(x)$. The quotient

$$q(x) = \frac{f(x)}{x-1}$$

has the property that

$$x^{n-1}q(\frac{1}{x}) = \frac{x^{n-1}f(\frac{1}{x})}{\frac{1}{x}-1} = \frac{-x^n f(\frac{1}{x})}{x-1} = \frac{f(x)}{x-1} = q(x) \ .$$

Therefore the same conclusion holds.

(3) In this case, $f(x)$ is itself in the desired form.

(4) Since $n = 2m$ is even and lower sign holds. In particular $a_m = a_{2m-m} = -a_m$, therefore $a_m = 0$ and

$$f(x) = (x^{2m} - 1) + a_1(x^{2m-1} - x) + a_2(x^{2m-2} - x^2) + \cdots \\ + a_{m-1}(x^{m+1} - x^{m-1}) .$$

Both 1 and $-1$ are roots of $f(x) = 0$. Therefore $(x^2 - 1)$ is a factor of $f(x)$. Using similar argument we verify that the quotient

$$q(x) = \frac{f(x)}{x^2 - 1}$$

also has the property

$$x^{n-2}q(\frac{1}{x}) = q(x) .$$

Thus the same conclusion holds.

5.4.4 EXAMPLE. *Solve* $x^5 - 1 = 0$.

SOLUTION: Clearly 1 is a root. For the quotient

$$\frac{x^5 - 1}{x - 1} = x^4 + x^3 + x^2 + x + 1$$

we have the equations

$$(x^2 + \frac{1}{x^2}) + (x + \frac{1}{x}) + 1 = 0$$

and

$$y^2 + y - 1 = 0 .$$

The roots of the last equation are $-\frac{1}{2}(1 + \sqrt{5})$ and $-\frac{1}{2}(1 - \sqrt{5})$. Thus we have the equations

$$x^2 + \frac{1}{2}(1 + \sqrt{5})x + 1 = 0$$
$$x^2 + \frac{1}{2}(1 - \sqrt{5})x + 1 = 0 .$$

Solving these we see the roots of the equation $x^5 - 1 = 0$ are 1, and $\frac{1}{4}(-1-$

$\sqrt{5} \pm i\sqrt{10-2\sqrt{5}}$), $\frac{1}{4}(-1+\sqrt{5} \pm i\sqrt{10+2\sqrt{5}})$ which are the five fifth roots of unity in radical form.

## EXERCISE 5D

1. Determine which of the following equations are reciprocal equations with symmetric or skew-symmetric coefficients.
   (a) $x^5 + x^4 - x^3 - x^2 + x + 1 = 0$.
   (b) $x^5 - x^4 + x^3 - x^2 + x - 1 = 0$.
   (c) $x^6 + x^5 + x^4 - x^3 - x^2 - x - 1 = 0$.
   (d) $x^6 - x^5 + x^4 - x^3 + x^2 - x + 1 = 0$.
   (e) $x^6 - x^5 + x^4 - x^3 - x^2 + x - 1 = 0$.
   (f) $x^6 + 3x^5 + x^4 - x^2 - 3x - 1 = 0$.
   (g) $x^n - 1 = 0$, $n \in \mathbf{N}$.
   (h) $x^n + 1 = 0$, $n \in \mathbf{N}$.
   (i) $(1 + ax + x^2)^n = 0$, $a \in \mathbf{R}$, $n \in \mathbf{N}$.
   (j) $(1 + x)^n = b(1 + x^n)$, $b \in \mathbf{R}$, $n \in \mathbf{N}$.
2. Solve the following equations.
   (a) $x^5 + 2x^4 + 3x^3 + 3x^2 + 2x + 1 = 0$.
   (b) $3x^5 - 10x^4 + 3x^3 - 3x^2 + 10x - 3 = 0$.
   (c) $x^5 - 2x^4 + 2x^3 - 2x^2 + 2x - 1 = 0$.
   (d) $x^6 + 4x^5 + 4x^4 - 4x^2 - 4x - 1 = 0$.
   (e) $x^7 - x^4 + x^3 - 1 = 0$.
3. The following equation, though not reciprocal, may be solved in a similar manner:
   $$6x^4 - 25x^3 + 12x^2 + 25x + 6 = 0 .$$
   Solve the equation.
4. Given $x^6 + ax^5 + bx^4 + cx^3 + bx^2 + ax + 1 = 0$, where $a$, $b$, and $c$ are real numbers. If after the transformation by putting $y = x + \frac{1}{x}$, a reciprocal equation in $y$ is obtained, express $b$ and $c$ in terms of $a$.
5. Find the real number $p$ if the product of two of the roots of $x^4 + px^3 + 3x^2 + px + 1 = 0$ is 2.
6. If the product of two of the roots of $x^4 + px^3 + qx^2 + px + 1 = 0$ is 2, prove that $4p^2 - 18q + 45 = 0$.

## CHAPTER SIX

# BOUNDS OF REAL ROOTS

Let given be an equation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0$$

where the coefficients have numerical real values. In our attempt to find the real roots of $f(x) = 0$, it would be very advantageous if we knew the range of values in which they might occur. To put it in another way, we wish to obtain for the search of the real roots of $f(x) = 0$ an upper bound $U$ so that a real number $s$ will not be a root if $s > U$, and a lower bound $L$ so that $s$ will not be a root if $s < L$. For some equations such bounds can be readily found. For example, if all coefficients are non-negative, then 0 is an upper bound; and 0 will be a lower bound if the signs of the coefficients are alternating.

## 6.1 The leading term

Let us study the numerical equation

$$f(x) = 2x^4 + 12x^3 - 36x^2 - 38x - 48 = 0 \ .$$

The polynomial $f(x)$ is the sum of its terms, $2x^4, 12x^3, -36x^2, -38x$ and $-48$. The terms as well as the polynomial $f(x)$ itself are all functions of one variable. Among the terms of $f(x)$, the leading term $2x^4$ stands out as the function with the fastest growth as shown in the table below:

| $x$ | $2x^4$ | $12x^3$ | $-36x^2$ | $-38x$ | $-48$ | $f(x)$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | −48 | −48 |
| 1 | 2 | 12 | −36 | −38 | −48 | −108 |
| 10 | 20000 | 12000 | −3600 | −380 | −48 | 27972 |
| 20 | 320000 | 96000 | −14400 | −760 | −48 | 400792 |

Initially at $x = 0, 1$ the term $2x^4$ is still inferior to some of the other terms, but it soon outgrows them by leaps and bounds thereafter. For this superiority of the leading term we shall say that the leading term is *predominant* for large values of $x$. Moreover, we see that as soon as $x > 10$, the term $2x^4$ and the polynomial $f(x)$ will have the same sign. Therefore 10 is an upper bound of the real roots of the equation $f(x) = 0$. Similarly we find that $-10$ is a lower bound. As the equation $f(x) = 0$ has only two real roots, 3 and $-8$ (see Example 5.2.3), we see that they do lie between these bounds.

Let us now consider an equation

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

with general coefficients and $a_n \neq 0$. We can write the polynomial $f(x)$ into

$$f(x) = a_n x^n \{1 + \frac{a_{n-1}}{a_n}\frac{1}{x} + \cdots + \frac{a_1}{a_n}\frac{1}{x^{n-1}} + \frac{a_0}{a_n}\frac{1}{x^n}\} .$$

When $x$ tends towards infinity, the expressions $1/x, 1/x^2, \ldots, 1/x^n$ will all tend towards zero. Therefore the expression within the brackets will tend towards 1. But this also means that the polynomial $f(x)$ and the leading term $a_n x^n$ would have about the same value when $c$ with a sufficiently large absolute value $|c|$ is assigned to the variable $x$. In other words, also for a polynomial with general coefficients the leading term is predominant for large enough absolute values of $x$.

We shall now use the predominance of the leading term to find a pair of bounds for the real roots of the equation $f(x) = 0$. For this purpose it is sufficient to find a positive number $K$ such that for all $|s| > K$,

$$|a_n s^n| > |a_{n-1}s^{n-1} + \cdots + a_1 s + a_0| .$$

Because the inequality would imply that

$$\begin{aligned}|f(s)| &= |a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0| \\ &\geq |a_n s^n| - |a_{n-1}s^{n-1} + \cdots + a_1 s + a_0| > 0 .\end{aligned}$$

Thus $f(x) \neq 0$ if $s > K$ or $s < -K$; and hence the real roots of $f(x) = 0$ can only occur in the closed interval $[-K, K]$. The following theorem yields one such value of $K$.

6.1.1 THEOREM. *Given a polynomial* $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ *with* $a_n \neq 0$. *Let* $k = \max\{|a_{n-1}|, |a_{n-2}|, \ldots, |a_0|\}$ *and* $K = \frac{k}{|a_n|} + 1$. *Then*

$$|a_n s^n| > |a_{n-1}s^{n-1} + \cdots + a_1 s + a_0|$$

*if* $|s| \geq K$.

PROOF: Let $k$ and $K$ have the values given in the theorem, and let $s$ be a real number. Now if $|s| \geq K$, then $\frac{k}{|s|-1} \leq |a_n|$. Therefore

$$\begin{aligned}
|a_{n-1}s^{n-1} + \cdots + a_1 s + a_0| &\leq |a_{n-1}s^{n-1}| + \cdots + |a_1 s| + |a_0| \\
&\leq k(|s|^{n-1} + \cdots + |s| + 1) \\
&= \frac{k}{|s|-1}(|s|^n - 1) \\
&\leq |a_n|(|s|^n - 1) \\
&< |a_n s^n| \,.
\end{aligned}$$

Hence $|a_n s^n| > |a_{n-1}s^{n-1} + \cdots + a_1 s + a_0|$.

Applying Theorem 6.1.1 to the equation $2x^4 + 12x^3 - 36x^2 - 38x - 48 = 0$, we get $k = 48$ and $K = 25$. Thus a pair of bounds 25 and $-25$ are easily found by this method. In comparison with the old pair 10 and $-10$, which we got for the same polynomial earlier, we find the new pair easier to be evaluated but inferior for our purpose.

## EXERCISE 6A

1. Using Theorem 6.1.1, find a pair of bounds for the real roots of each of the following equations.
   (a) $2x^4 + 12x^3 + 17x^2 + 14x + 6 = 0$.
   (b) $2x^4 + 3x^3 + 9x^2 - 5x - 6 = 0$.
   (c) $2x^4 + x^3 - 5x^2 - 7x - 6 = 0$.
   (d) $2x^4 - 5x^3 + x^2 - x + 6 = 0$.

2. Find real numbers $a$, $b$, and $c$ such that

$$x^5+x^4-100x^3-119x^2+x-132 = x^3(x^2-a)+x^2(x^2-b)+(x-c)(x+12)\,.$$

Hence, find the smallest possible integral upper bound for the roots of the equation $x^5 + x^4 - 100x^3 - 119x^2 + x - 132 = 0$. Compare this bound with that given by Theorem 6.1.1. Which is better?

3. Use the technique employed in Question 2 to find a better upper bound for the real roots of each equation of (c) and (d) in Question 1.

(Note: There may not be a unique answer for each equation and it depends on how you group the terms together.)

4. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ and $A = \max\{|a_{n-1}|, |a_{n-2}|, \ldots, |a_0|\}$.

   (a) Show that all real roots of $f(x) = 0$ are less than or equal to $1 + A$.

   (b) By considering $f(\frac{1}{x}) = 0$, show that all real positive roots of $f(x) = 0$ are greater than or equal to $\frac{1}{1+B}$, where $B = \max\left\{\frac{1}{|a_0|}, |\frac{a_{n-1}}{a_0}|, \ldots, |\frac{a_1}{a_0}|\right\}$.

   (c) By considering $f(-x) = 0$, show that the negative roots of $f(x) = 0$ are greater than or equal to $-(1 + A)$.

5. By using the results in Question 4, find upper bounds and lower bounds for the positive and negative roots of $x^5 - 2x^4 - 5x^3 - 8x^2 - 7x + 3 = 0$.

## 6.2 The constant term

Similar to the leading term which has been found to be predominant for 'large' values of the unknown $x$, the constant term becomes predominant for 'small' values. To see this, we write $f(x) = a_nx^n + \cdots + a_1x + a_0$ with $a_0 \neq 0$ into

$$f(x) = a_0\{1 + \frac{a}{a_0}x + \cdots + \frac{a_{n-1}}{a_0}x^{n-1} + \frac{a_n}{a_0}x^n\}$$

when $x$ tends towards zero, the expressions $x, x^2, \ldots, x^n$ all tend towards zero. Therefore, when $0 < |x|$ is sufficiently small, the polynomial $f(x)$ will hardly be distinguishable from the constant $a_0$. This predominance of the constant term can be used to find a lower bound

for the positive roots and an upper bound of the negative roots of $f(x) = 0$. For this purpose we prove the following theorem which is parallel to Theorem 6.1.1.

6.2.1 THEOREM. *Given a polynomial* $f(x) = a_0 + a_1x + \cdots + a_nx^n$ *with* $a_0 \neq 0$ *and* $a_n \neq 0$. *Let* $g = \max\{|a_1|, |a_2|, \ldots, |a_n|\}$ *and* $H = |a_0|/(|a_0|+g)$. *Then*

$$|a_0| > |a_1s + a_2s^2 + \cdots + a_ns^n| \,.$$

*if* $|s| \leq H$.

PROOF: Consider the polynomial $h(y) = y^n f(\frac{1}{y})$. Then

$$h(y) = a_0y^n + a_1y^{n-1} + \cdots + a_{n-1}y + a_n \,.$$

Between the given polynomial $f(x)$ and its 'transform' $h(y)$, there is a formal relation: if $r \neq 0$ is a root of $f(x)$ then $1/r$ is a root of $h(y)$, and vice versa. Moreover if we apply Theorem 6.1.1 to the polynomial $h(y)$ we find the value $k$ of $h(y)$ identical to $g$ and the value $K$ of $h(y)$ identical to $1/H$ because

$$\frac{1}{H} = \frac{|a_0| + g}{|a_0|} = \frac{g}{|a_0|} + 1 = \frac{k}{|a_0|} + 1 = K \,.$$

Therefore if $|s| \leq H$, then $|1/s| \geq K$ and by 6.1.1, we get

$$|a_0\frac{1}{s^n}| > |a_1\frac{1}{s^{n-1}} + a_2\frac{1}{s^{n-2}} + \cdots + a_n| \,.$$

Hence $|a_0| > |a_1s + a_2s^2 + \cdots + a_ns^n|$.

The theorem gives $H$ as a lower bound of the positive roots and $-H$ as an upper bound of the negative roots. We remark here that the value of $H = |a_0|/(|a_0| + g)$ lies between 0 and 1. Therefore in general, the theorem only provides rather poor values for these bounds. In some cases, it may even be quite useless. Take, for example, the equation $x^2 - 50x + 5000 = (x - 100)(x + 50) = 0$ which has one positive root 100 and one negative root $-50$. But the lower bound of the positive roots provided by the theorem is less than 1 and the upper bound greater than $-1$. Therefore the theorem will only give us some useful information for a search of roots lying between $-1$ and 1. Nevertheless, inspite of the fact that for individual polynomials,

Theorems 6.1.1 and 6.2.1 may only provide us with rather crude estimates of the bounds, the existence of such bounds is of immense importance for our study of certain analytic properties of polynomial functions. The following corollary will be useful in the next three chapters.

COROLLARY 6.2.2. *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial such that $a_n \neq 0$ and $a_0 \neq 0$. Then for sufficiently large positive $c$, $f(c)$ and $a_n$ will have the same sign; and for sufficiently small positive $h$, $f(h)$ and $a_0$ will have the same sign.*

EXERCISE 6B

1. Use Theorem 6.2.1 to find a lower bound for the positive real roots and an upper bound for the negative real roots of each of the following equations.
   (a) $2x^4 + 12x^3 + 17x^2 + 14x + 6 = 0$.
   (b) $2x^4 + 3x^3 + 9x^2 - 5x - 6 = 0$.
   (c) $2x^4 + x^3 - 5x^2 - 7x - 6 = 0$.
   (d) $2x^4 - 5x^3 + x^2 - x + 6 = 0$.
2. Combining the results found in Question 1 of Exercise 6A, write down a pair of bounds for the positive real roots and a pair of bounds for the negative roots of the equations in Question 1.
3. Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ and $r$ is a real root of $f(x) = 0$. Show that $|r| > H$, where $H = \frac{|a_0|}{|a_0| + g}$ and $g = \max\{|a_1|, |a_2|, \ldots, |a_n|\}$.

### 6.3 Other bounds of real roots

The chief concern of the theory of equation is the evaluation of the roots of a given equation. On appearance the two theorems of the last sections seem to provide us with information on the location of the roots because together they state that real roots of the equation $f(x) = 0$ with non-vanishing constant term can only be found in the

intervals $[H, K]$ and $[-K, -H]$. However in some cases the values of $K$ and $H$ may be too crude to be really useful. Take, for example, $f(x) = x^3 + 20x^2 + 75x - 1000$. In this case $K = 1001$. Let us also inspect the following table of values.

| $x$ | $x^3$ | $20x^2$ | $75x$ | $-1000$ | $f(x)$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | $-1000$ | $-1000$ |
| 1 | 1 | 20 | 75 | $-1000$ | $-904$ |
| 5 | 125 | 500 | 375 | $-1000$ | 0 |
| 10 | 1000 | 2000 | 750 | $-1000$ | 2750 |

We see that $f(x) = 0$ has no positive root greater than 5. Therefore, as an upper bound of roots, $K = 1001$ is too large to be useful. If we read the proof of Theorem 6.1.1 more carefully, we shall discover that we have failed to take into consideration the signs and the exponents of the terms $a_i x^i$, and consequently overestimated $K$.

In this section, we shall make the necessary remedy and obtain better values for the bounds of real roots. We observe that if the coefficients of $f(x) = 0$ are all non-negative, then the equation would have no positive root and 0 could serve as a bound. Therefore we need only consider polynomials in which some coefficients are negative.

6.3.1 THEOREM. *Given a polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ in which some coefficients are negative and the leading coefficient equals 1. Let $a_r$ be the first negative coefficient (i.e. $a_r < 0$ and $a_{r+1} \geq 0, a_{r+2} \geq 0, \ldots, a_{n-1} \geq 0$) and let $-G$ be the least of all coefficients (i.e. $-G \leq a_i$ for $i = 0, 1, \ldots, n-1$). Then $f(s) > 0$ for any real number $s \geq 1 + \sqrt[n-r]{G}$.*

PROOF: It follows from the definition of $G$ and $r$ that

$$\begin{aligned} f(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \\ &\geq s^n + a_r s^r + \cdots + a_1 s + a_0 \\ &\geq s^n - G(s^r + \cdots + s + 1) \\ &= s^n - G\frac{s^{r+1} - 1}{s - 1} . \end{aligned}$$

Therefore it suffices to show that

$$\text{if} \quad s \geq 1 + \sqrt[n-r]{G} \quad \text{then} \quad G(s^{r+1} - 1) < s^n(s-1).$$

Let $s \geq 1 + \sqrt[n-r]{G}$. Since $f(x)$ has some negative coefficient we conclude that $0 < G \leq (s-1)^{n-r}$ and $1 < s$. Then

$$\begin{aligned} G(s^{r+1} - 1) < Gs^{r+1} \leq s^{r+1}(s-1)^{n-r} &= s^{r+1}(s-1)(s-1)^{n-r-1} \\ &< s^{r+1}(s-1)s^{n-r-1} = s^n(s-1) . \end{aligned}$$

The proof is now complete.

The expression $1 + \sqrt[n-r]{G}$ given in the above theorem in terms of the coefficients is therefore an upper bound of all positive roots of the equation $x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$ with some negative coefficients. For the equation $x^3 + 20x^3 + 75x - 1000 = 0$, considered earlier in this section, we get an upper bound $1 + \sqrt[3-0]{1000} = 11$ which is far better than the value 1001 provided by Theorem 6.1.1. As for the equation $2x^4 + 12x^3 - 36x^2 - 38x - 48 = 0$ which we studied earlier in Section 6.1, we get $1 + \sqrt[4-2]{24}$. Thus 6 can be taken as an upper of the positive roots of the equation which is also better than the previous values of 10 and 25. In fact 6 is quite close to the only positive roots 3 of the equation.

The formulation of Theorem 6.3.1 may seem somewhat cumbersome at first sight. The following examples will show that it is quite handy for applications.

6.3.2 EXAMPLE. *The equation $x^4 - 5x^3 + 40x^2 - 8x + 24 = 0$ has 9 as an upper bound of its real roots.*

PROOF: Using the notation of Theorem 6.3.1 we have $n = 4$, $r = 3$, $G = 8$. Therefore $1 + \sqrt[n-r]{G} = 9$. By Theorem 6.3.1, $f(s) > 0$ for all $s \geq 9$. Therefore all roots of $f(x) = 0$ must be less than 9, i.e. 9 is an upper bound of the real roots.

6.3.3 EXAMPLE. *Find an upper bound of the real roots of the equation $3x^5 + 9x^4 + 3x^3 - 24x^2 - 153x + 54 = 0$ .*

SOLUTION: Divide the equation by 3 to get

$$x^5 + 3x^4 + x^3 - 8x^2 - 51x + 18 = 0 .$$

Therefore $r = 2$, $n - r = 3$, $G = 51$. Thus $1 + \sqrt[3]{51}$ or 5 is an upper bound of the real roots.

By a suitable transformation of the equation $f(x) = 0$, Theorem 6.3.1 can provide us with a lower bound of roots. Let

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a polynomial with real coefficients. Consider the transformed polynomial

$$g(y) = (-1)^n f(-y) = (-1)^n \{(-y)^n + a_{n-1}(-y)^{n-1} + \cdots + a_1(-y) + a_0\} .$$

Then the relation between the equations $f(x) = 0$ and $g(y) = 0$ is such that a real number $-r$ is a root of the former if and only if $r$ is a root of the latter and vice versa. Therefore the negative of an upper bound of the positive roots of $g(y) = 0$ will be a lower bound of the negative roots of $f(x) = 0$.

Because of the superiority of Theorem 6.3.1 over Theorem 6.1.1, the lower bounds obtained by the method above are usually better than those obtained by the previous method. Take again the equation $x^3 + 20x^2 + 75x - 1000 = 0$ which has only one real root 5. The lower bound by the previous method is $-1001$. Using the present method, we obtain the equation $y^3 - 20y^2 + 75y + 1000 = 0$ which has an upper bound 21. Therefore as a lower bound for the given equation in $x$, $-21$ is far better than $-1001$.

6.3.4 EXAMPLE. *The equation $x^4 - 5x^3 + 40x^2 - 8x + 24 = 0$ has no negative root.*

PROOF: Substituting $-y$ for $x$ in the equation, we get $y^4 + 5y^3 + 40y^2 + 8y + 24 = 0$. This equation has no negative coefficient and hence has no positive root. Therefore the given equation has no negative root.

6.3.5 EXAMPLE. *Find a lower bound of the roots of the equation $3x^5 + 9x^4 + 3x^3 - 24x^2 - 153x + 54 = 0$ of Example 6.3.3.*

SOLUTION: Substituting $-y$ for $x$ and dividing the resulting equation by $-3$, we get $y^5 - 3y^4 + y^3 + 8y^2 - 51y - 18 = 0$. By Theorem 6.3.1, we get $1+51 = 52$ as an upper bound of the positive roots of the last equation in $y$. Therefore $-52$ is a lower bound of the negative roots of the given equation in $x$.

When we use Theorem 6.3.1 in conjunctions with other suitable transformations we can obtain other bounds of roots. Let

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

again be a polynomial with real coefficients. Consider the transformed polynomial

$$h(z) = z^n f(\frac{1}{z}) \ .$$

Then the relation between the equations $f(x) = 0$ and $h(z) = 0$ is such that a non-zero real number $1/r$ is a root of $f(x) = 0$ if and only if $r$ is a root of $h(z) = 0$ and vice versa. Therefore, if $K > 0$ is an upper bound of the positive roots of $h(z) = 0$ then $\frac{1}{K}$ is a lower bound of the positive roots of $f(x) = 0$.

6.3.6 EXAMPLE. *All real roots of the equation $x^4-5x^3+40x^2-8x+24 = 0$ (Examples 6.3.2 and 6.3.4) belong to the open interval $(\frac{3}{4}, 9)$ of the real line.*

PROOF: By Example 6.3.4 and the fact that the constant term of the equation is non-zero, the equation can only have positive roots. By Example 6.3.2, all roots of the equation are less than 9. To find a lower bound, we substitute $1/z$ for $x$ and multiply the resulting equation by $z^4$ to get

$$1 - 5z + 40z^2 - 8z^3 + 24z^4 = 0 \ .$$

Divide this equation by 24 to get

$$z^4 - \frac{1}{3}z^3 + \frac{5}{3}z^2 - \frac{5}{24}z + \frac{1}{24} = 0 \ .$$

Now $n = 4$, $r = 3$ and $G = \frac{1}{3}$. Therefore the roots of this equation must be less than $1 + \frac{1}{3} = \frac{4}{3}$. Thus the positive roots of the original equation must be greater than $\frac{3}{4}$. Therefore they belong to the interval $(\frac{3}{4}, 9)$.

## EXERCISE 6C

1. Using Theorem 6.3.1 find an upper bound for the real roots of the following equations.
   (a) $x^4 - 4x^3 + 2x^2 + x + 6 = 0$.
   (b) $x^4 - 6x^2 - 7x - 6 = 0$.
   (c) $x^4 + 2x^3 - 4x^2 - 5x - 6 = 0$.
   (d) $2x^4 + x^3 - 5x^2 - 7x - 6 = 0$.
   (e) $2x^4 + 3x^3 + 9x^2 - 5x - 6 = 0$.
2. By putting $y = -x$, find a lower bound for the real roots of each equation in Question 1.
3. By using the transformation $y = \frac{1}{x}$, find a pair of bounds for the positive real roots of each equation in Question 1.
4. Find the upper bounds and lower bounds of the positive real roots and the negative real roots of $x^5 + 2x^4 - 5x^3 + 8x^2 - 7x - 3 = 0$ by Theorem 6.3.1.

Consider $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ of $\mathbf{R}[x]$ in Questions 5 to 7.

5. (a) Let $s$ be a positive real number. Prove that the roots $r$ of $f(x) = 0$ satisfy

$$|r| \leq \max\{s, |a_{n-1}| + |a_{n-2}|s^{-1} + \cdots + |a_0|s^{1-n}\}$$

   and hence show that

$$|r| \leq s \quad \text{if} \quad s \geq |a_{n-1}| + |a_{n-2}|s^{-1} + \cdots + |a_0|s^{1-n} \, .$$

   (b) Show that $|r| \leq \max\{1, |a_{n-1}| + |a_{n-2}| + \cdots + |a_0|\}$ for any root $r$ of $f(x) = 0$.

6. (a) If $s = |a_{n-1}| + |a_{n-2}|^{\frac{1}{2}} + \cdots + |a_0|^{\frac{1}{n}}$, prove that $s^{1-i} \leq |a_{n-i}|^{\frac{1}{i}} \cdot |a_{n-i}|^{-1}$, for $i = 1, \ldots, n$.

   (b) By (a) and Question 5, show that any root $r$ of $f(x) = 0$ satisfies $|r| \leq |a_{n-1}| + |a_{n-2}|^{\frac{1}{2}} + \cdots + |a_0|^{\frac{1}{n}}$.

7. (a) If $s = |a_{n-1}| + \frac{|a_{n-2}|}{|a_{n-1}|} + \cdots + \frac{|a_0|}{|a_1|}$ with $a_1 \ldots a_{n-1} \neq 0$, prove that $s^{i-1} \geq |a_{n-i+1}|$ for $i = 1, \ldots, n$.

   (b) By (a) and Question 5, show that any root $r$ of $f(x) = 0$ satisfies $|r| \leq |a_{n-1}| + \frac{|a_{n-2}|}{|a_{n-1}|} + \cdots + \frac{|a_0|}{|a_1|}$.

CHAPTER SEVEN

# THE DERIVATIVE

Up to the last chapter, only purely algebraic properties of polynomials are used in our study of equations. Beginning with this chapter, we shall put more emphasis on the functional aspect of the polynomial and examine in detail the change of the value of a polynomial corresponding to a minute increase or diminution of the variable. This will lead us to the discovery of certain basic analytic properties of polynomials such as continuity and differentiability which are usually within the purview of calculus.

## 7.1 Differentiation

Readers who are familiar with the techniques of elementary calculus will recall that for a certain type of real valued functions $f(x)$ of one real variable, at every point $c$ of the domain the limit

$$\lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

exists and is called the *derivative of $f(x)$ at the point $c$* and denoted by $f'(c)$. The function that takes $c$ to $f'(c)$ for every $c$ is itself call the *derivative* of $f(x)$. In geometric terms, $f'(c)$ is the slope of the tangent to the curve $y = f(x)$ at the point $(c, f(c))$. Alternatively the derivative $f'(x)$ can be interpreted as the rate of change of the varying quantity $f(x)$; for example, if $v(t)$ represents the velocity of a moving body at time $t$, then $v'(t)$, being the rate of change of velocity, is the acceleration of the moving body at time $t$. The type of functions that possess a derivative include the polynomials and other elementary functions as well as many other functions. Isaac Newton (1642–1727) and Gottfried Wilhelm Leibniz (1646–1716) independently made use of the derivative in their separate discovery of calculus which had tremendous impact on the development of mathematics and science.

In this chapter, we study the analytic properties of the derivative of a polynomial function and use them in our study of polynomials and equations. Instead of borrowing the definition of derivative from calculus, we shall start afresh by a purely algebraic approach to arrive at the same definition without using limit and convergence.

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial with real coefficients. Then $f(x) : \mathbf{R} \to \mathbf{R}$ is a real-valued function in one variable. For a fixed point $c$ of the domain, every point in a neighbourhood of $c$ can be represented by a number $c + h$. If we regard $h$ as a variable quantity, then $c+h$ becomes a varying point of the neighbourhood. We proceed to investigate the relation between the fixed functional value

$$f(c) = a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0$$

and the varying functional value

$$f(c+h) = a_n(c+h)^n + a_{n-1}(c+h)^{n-1} + \cdots + a_1(c+h) + a_0$$

in terms of the variable quantity $h$. After expanding the binomials on the right-hand side of the last equality and collecting like terms that have the same exponents of $h$, we obtain a polynomial in the variable $h$:

$$f(c+h) = D_0 + D_1 h + D_2 h^2 + \cdots + D_n h^n \ .$$

The coefficients

$$\begin{aligned}
D_0 &= a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0 \\
D_1 &= n a_n c^{n-1} + (n-1)a_{n-1}c^{n-2} + \cdots + 2c_2 a + a_1 \\
D_2 &= \frac{1}{2!}\{n(n-1)a_n c^{n-2} + (n-1)(n-2)a_{n-1}c^{n-3} + \cdots + 2c_2\} \\
&\cdots\cdots\cdots\cdots\cdots \\
D_n &= \frac{1}{n!}\{n(n-1)(n-2)\cdots 2 \cdot 1 \cdot a_n\}
\end{aligned}$$

are all polynomial expressions in the fixed quantity $c$.

Clearly the first coefficient $D_0$ is identical to $f(c)$:

$$D_0 = a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0 = f(c)$$

which is a familiar expression. The second coefficient

$$D_1 = na_nc^{n-1} + (n-1)a_{n-1}c^{n-2} + \cdots + 2a_2c + a_1$$

of the linear term $D_1h$ can be obtained by a simple transformation of $D_0$ in which each of the $n+1$ terms $a_rc^r$ of $D_0$ becomes a term $ra_rc^{r-1}$ of $D_1$. Thus $D_1$ has $n$ terms because the last term $a_0$ of $D_0$ becomes 0 in $D_1$. This real number $D_1$, so obtained by purely algebraic means, is in fact the derivative of the function $f(x)$ at $x = c$. To justify this statement, we must compare $D_1$ with the analytic definition of derivative:

$$f'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h} .$$

Substituting for $f(c+h)$ the polynomial expression in $h$ and taking into consideration that $D_0 = f(c)$, we get

$$\begin{aligned} f'(c) &= \lim_{h\to 0} \frac{1}{h}\{f(c+h) - f(c)\} \\ &= \lim_{h\to 0}\{D_1 + D_2h + \cdots + D_nh^{n-1}\} \\ &= D_1 . \end{aligned}$$

Therefore the real number $D_1$ is the derivative of $f(x)$ at $x = c$ which shall be denoted by $f'(c)$ as defined in calculus. Consequently the polynomial

$$f'(x) = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2x + a_1$$

is the derivative of $f(x)$.

7.1.1 DEFINITION. Given a polynomial

$$f(x) = a_nx^n + a_{n-1}x^n + \cdots + a_1x + a_0$$

of degree $n$, the polynomial

$$f'(x) = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2x + a_1$$

of degree $n-1$ is called the *derivative* of $f(x)$, and for any real number $c$, the real number

$$f'(c) = na_nc^{n-1} + (n-1)a_{n-1}c^{n-2} + \cdots + 2a_2c + a_1$$

is called the *derivative of* $f(x)$ *at* $x = c$.

We take note that to obtain $f'(x)$ from $f(x)$ we

(i) *multiply each term of* $f(x)$ *by its exponent:*

$$na_nx^n, (n-1)a_{n-1}x^{n-1}, \dots, 2a_2x^2, 1a_1x, 0a_0 \; .$$

(ii) *diminish the exponent of each monomial by* 1*:*

$$na_nx^{n-1}, (n-1)a_{n-1}x^{n-2}, \dots, 2a_2x, a_1, 0 \; .$$

(iii) *add up to get*

$$f'(x) = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2x + a_1 \; .$$

For instance, given

$$f(x) = 3x^3 - 7x + 2 \quad \text{and} \quad g(x) = 5x^9 - 2x^7 + x^2$$

we follow the steps (i), (ii) and (iii) to obtain their derivatives

$$f'(x) = 9x^2 - 7 \quad \text{and} \quad g'(x) = 45x^8 - 14x^6 + 2x \; .$$

From now on the apostroph, when used in conjunction with a polynomial symbol, is reserved exclusively for the notation of the derivative. Thus $f'(x)$ or $f(x)'$ can only have the meaning of the derivative of $f(x)$. Furthermore, the algebraic operation of forming the derivative is called the *differentiation* which is a mapping $f(x) \to f'(x)$ of $\mathbf{R}[x]$ into itself. Two formal algebraic properties of differentiation are formulated in the following theorem.

7.1.2 THEOREM. *Let* $f(x)$ *and* $g(x)$ *be polynomials of* $\mathbf{R}[x]$. *Then the derivative of their sum is* $f'(x) + g'(x)$ *and the derivative of their product is* $f'(x)g(x) + f(x)g'(x)$, *i.e.*

$$(f(x) + g(x))' = f'(x) + g'(x) \quad \textit{and} \quad (f(x)g(x))' = f'(x)g(x) + f(x)g'(x) \; .$$

PROOF: The statement concerning the sum is obviously true. For the product we observe that since the sum rule is true and every polynomial is just a sum of monomials, it suffices to prove the product rule for monomials. Let $f(x) = ax^n$ and $g(x) = bx^m$. Then $f(x)g(x) = abx^{n+m}$. There-

fore $(f(x)g(x))' = (n+m)abx^{n+m-1} = (nax^{n-1})(bx^m) + (ax^n)(mx^{m-1}) = f'(x)g(x) + f(x)g'(x)$.

We have called the operation of forming the derivative differentiation. Thus given a polynomial

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

we differentiate to get its derivative

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1 \ .$$

Similarly, we differentiate $f'(x)$ to get its derivative

$$(f'(x))' = n(n-1)x^{n-2} + (n-1)(n-2)a_{n-1}x^{n-3} + \cdots + 2a_2$$

which is by definition the derivative of the derivative of $f(x)$, conveniently denoted by $f''(x)$. In turn, $f''(x)$ can be differentiated to yield $f'''(x)$, etc. This leads us to introduce the following recursive definition and notation.

7.1.3 DEFINITION. The derivative $f'(x)$ of a polynomial $f(x)$ is also called the *first derivative* of $f(x)$ and may be denoted by $f^{(1)}(x)$. The $k-th$ *derivative* $f^{(k)}(x)$ of $f(x)$ is defined as the derivative of the $(k-1)-th$ derivative of $f(x)$. Thus $f^{(k)}(x) = (f^{(k-1)}(x))'$.

Usually the first, second and third derivatives of $f(x)$ are also denoted by $f'(x), f''(x)$ and $f'''(x)$. For higher derivatives the index notation is preferred. Since a differentiation diminishes the degree of a polynomial by 1, if $f(x)$ has a degree $n$, then $\deg f^{(k)}(x) = n-k$. In particular $f^{(n)}(x)$ is the non-zero constant $(n!)a_n$ and $f^{(n+1)}(x) = 0$. Occasionally $f(x)$ itself is called the $0-th$ *derivative*: $f^{(0)}(x) = f(x)$.

## EXERCISE 7A

In what follows, all the polynomials are in $\mathbf{R}[x]$.

1. A polynomial $f(x)$ of degree $n > 1$ has the property that $f(\alpha) = 0$ and $f'(\alpha) = 0$ for some real number $\alpha$. Prove that
$$f(x) = (x - \alpha)^2 g(x)$$
for some polynomial $g(x)$ of degree $n - 2$.

2. If $n$ is a positive integer and $f(x)$ is a polynomial, let $g(x) = [f(x)]^n$. Show that $g'(x) = n[f(x)]^{n-1} f'(x)$ without using the chain rule for differentiation.

3. Let $f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$. By using the binomial theorem, prove that
$$f(x+h) = f(x) + f'(x) \cdot h + f''(x) \cdot \frac{h^2}{2!} + f'''(x) \cdot \frac{h^3}{3!}$$
where $h$ is any real number.

4. Prove that the remainder on dividing $f(x)$ by $(x - a)^2$ is
$$f'(a)(x - a) + f(a).$$

5. Let $f(x)$ be any polynomial and $a$ is a real number. Define a linear polynomial $g(x)$ by $g(x) = f'(a)(x-a) + f(a)$. Prove that if $f(x) = x^n$, for any positive integer $n$, then $f(x) - g(x)$ is divisible by $(x - a)^2$. Can the result be extended to any polynomial in $\mathbf{R}[x]$? Justify your answer.

6. Let $f_0(x) = 1$, and for $n \geq 1$ define
$$f_n(x) = \frac{x(x-n)^{n-1}}{n!}.$$
Show that $f_n{}'(x) = f_{n-1}(x - 1)$ for $n \geq 1$ and deduce that
$$f_n(0) = f_n{}'(1) = f_n{}''(2) = \cdots = f_n{}^{(n-1)}(n-1) = 0, \quad \text{and}$$
$$f_n{}^{(n)}(n) = 1.$$

7. (a) If $g(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, find a polynomial $f(x)$ such that $g(x) = f'(x)$.

(b) (i) Find a polynomial $f(x)$ of degree $n$ such that $f'(x) = 0$ has $n-1$ real roots.

(ii) Find a polynomial $f(x)$ of odd degree $n$ such that $f'(x) = 0$ has no real root. How about if $n$ is even?

8. If $f(x)$ is a polynomial such that $f(0) = 0$ and $(x+2)f'(x) - 2f(x) + 2 = 0$. Find $f(x)$.

   [Hint: Write $f(x) = a_n x^n + g(x)$ where $g(x)$ is a polynomial of degree $< n$.]

9. (a) Prove the converse of Question 1, that is, if a polynomial $f(x)$ of degree $n > 1$ has the property that $f(x) = (x-\alpha)^2 g(x)$, for some real number $\alpha$ and polynomial $g(x)$, prove that $f(\alpha) = f'(\alpha) = 0$.

   (b) Let $f(x) = x^5 - x^3 + 4x^2 - 3x + 2$. Find HCF$(f(x), f'(x))$ and hence solve $f(x) = 0$.

10. For any $n+1$ real numbers $a_0, a_1, \cdots, a_n$ and any real value $x = x_0$, prove that there is a real polynomial $f(x)$ of degree $n$ such that

$$f^{(i)}(x_0) = a_i \qquad i = 0, 1, \cdots, n$$

where $f^{(0)}(x_0) = f(x_0)$.

11. Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be the $n$ roots of $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. We write $s_1 = \sum_{i=1}^{n} \alpha_i$, $s_2 = \sum_{i=1}^{n} \alpha_i{}^2$, and in general, $s_k = \sum_{i=1}^{n} \alpha_i{}^k$ for positive integer $k$.

    (a) Show that $f'(x) = \sum_{i=1}^{n} \frac{f(x)}{x-\alpha_i}$.

    (b) Show that $f'(x) = nx^{n-1} + (s_1 + na_{n-1})x^{n-2} + (s_2 + a_{n-1}s_1 + na_{n-2})x^{n-3} + \cdots + (s_k + a_{n-1}s_{k-1} + a_{n-2}s_{k-2} + \cdots + a_{n-k+1}s_1 + na_{n-k})x^{n-k-1} + \cdots + a_1$.

    (c) Prove that for $k = 1, 2, \cdots, n-1$,

$$s_k + a_{n-1}s_{k-1} + a_{n-2}s_{k-2} + \cdots + ka_{n-k} = 0\ .$$

    (d) By considering the equation $x^{k-n}f(x) = 0$, prove that

$$s_k + a_{n-1}s_{k-1} + \cdots + a_0 s_{k-n} = 0$$

    for positive integer $k \geq n$.

(The results in (c) and (d) are called Newton's formulae which enable us to express any $s_k$ in terms of $a_1, \cdots, a_{n-1}$.)

12. Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be roots of $x^n + nax - b = 0$, where $n$ is a positive integer.

    (a) Show that

    $$\prod_{i<j}(\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} \cdot n^n \cdot \prod_{i=1}^{n}(\alpha_i{}^{n-1} + a).$$

    [Hint : Consider two different expressions for the derivative of $x^n + nax - b$.]

    (b) Show that $\alpha_i{}^{n-1}$, $i = 1, 2, \cdots, n$ are the roots of the equation $x(x + na)^{n-1} - b^{n-1} = 0$.

    (c) Hence, show that

    $$\prod_{i<j}(\alpha_i - \alpha_j)^2 = (-1)^{\frac{(n-1)(n+2)}{2}} \cdot n^n \cdot \{(n-1)^{n-1}a^n + b^{n-1}\}.$$

### 7.2 Taylor's formula

Recall that at the beginning of the last section we set out to investigate the relation between the functional value $f(c)$ of $f(x)$ at $x = c$ and the functional value $f(c+h)$ at a neighbouring point of $c$. This led us to the very important expression

$$f(c+h) = D_0 + D_1 h + \cdots + D_n h^n$$

which is a polynomial in the variable $h$. In the last section we have identified $D_0 = f(c)$ and $D_1 = f'(c)$. Using the higher derivatives of $f(x)$ we have no difficulty in identifying the remaining coefficients:

$$D_k = \frac{1}{k!} f^{(k)}(c) .$$

Rewriting the coefficients $D_k$ throughout, we get

$$f(c+h) = f(c) + f'(c)h + \frac{f''(c)}{2!}h^2 + \cdots + \frac{f^{(k)}(c)}{k!}h^k + \cdots + \frac{f^{(n)}(c)}{n!}h^n$$

which is known as *Taylor's expansion* of $f(c+h)$. The expansion first

appeared in *Methodus Incrementorum* by Brook Taylor (1685–1731) though other mathematicians, e.g. Isaac Newton, have used such a device before him.

Treating $c$ as a variable point and replacing it by $x$, we may formulate the above expansion as

7.2.1 TAYLOR'S FORMULA. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$. Then*

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \cdots + \frac{f^{(k)}(x)}{k!}h^k + \cdots + \frac{f^{(n)}(x)}{n!}h^n.$$

For any fixed value of $x$ Taylor's formula expands the value $f(x+h)$ into a polynomial expression in $h$. We shall see in the subsequent sections and chapters that the formula is an indispensable tool in the study of the local behaviour of the function $f(x)$ in the neighbourhood of any given point. Finally, we remark that in addition to polynomials there is a very large class of real-valued functions for which a similar Taylor's formula holds.

## EXERCISE 7B

1. Let $f(x)$ be a polynomial of degree $n$ in $\mathbf{R}[x]$, and $a$ be any real number. Show that

$$f(x) = f(a) + f'(a)\cdot(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

   Hence express $x^3 - x^2 + 2x + 2$ as a polynomial in $x - 1$.

2. Consider $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ of $\mathbf{R}[x]$ where $a_n \neq 0$. Suppose for some real number $a$, $f(x) \geq 0$, $f'(a) \geq 0, \cdots, f^{(n)}(a) \geq 0$; prove that $f(x)$ has no root greater than $a$.
3. Let $f(x)$ be a monic real polynomial of degree $n > 1$, and $f(x)$ has $n$ real roots. Prove that the real number $b$ is an upper bound for the roots of $f(x)$ if and only if $f(b) \geq 0$, $f'(b) \geq 0, \cdots, f^{(n)}(b) \geq 0$.
4. Let $f(x)$ be a real polynomial of degree $n > 1$. Suppose $\alpha_1, \alpha_2, \cdots, \alpha_n$

are the roots of $f(x)$ and $c \neq \alpha_i$ for all $i$, prove by the Taylor's formula that

$$\sum_{i=1}^{n} \frac{1}{\alpha_i - c} = -\frac{f'(c)}{f(c)}.$$

## 7.3 Multiple roots

We recall that a real number $r$ is a *root of multiplicity* $k \geq 1$ of a polynomial equation $f(x) = 0$ if $f(x)$ is divisible by $(x-r)^k$ but not by $(x-r)^{k+1}$. A root of multiplicity 1 is called a *simple root* and any root of multiplicity $k > 1$ is called a *multiple root.* Multiple roots which are complex numbers are defined similarly. A root of multiplicity $k$ is also called a *k-fold root.* Thus it follows from the product rule of degrees that a polynomial $f(x)$ of $\mathbf{R}[x]$ of degree $n \geq 1$ has exactly $n$ roots if each multiple root is counted by its multiplicity (i.e. a $k$-fold root is counted as $k$ roots). For example, if $f(x) = 7(x-4)(x-3)^2(x+2)^3$, then $f(x) = 0$ has six roots: 4 counted once, 3 counted twice and $-2$ counted thrice.

As a first application of the results of the last section, we shall establish a relation between multiple roots of $f(x)$ and the derivatives of $f(x)$. To begin, we write down Taylor's formula

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots .$$

Substituting $c$ for $x$ and $x-c$ for $h$ in the above, we obtain

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^n(c)}{n!}(x-c)^n$$

which is an expression of $f(x)$ as a polynomial in the new variable $(x-c)$. Treating $f(x)$ as such and dividing it by $(x-c)$, we get

$$f(x) = (x-c)q_1(x) + f(c)$$

where the remainder is the constant $f(c)$ and the quotient,

$$q_1(x) = f'(c) + \frac{f''(c)}{2!}(x-c) + \cdots + \frac{f^n(c)}{n!}(x-c)^{n-1}$$

obtainable from the above expression, is expressed as a polynomial in $(x-c)$ of degree $n-1$. Similarly divisions by $(x-c)^k$ and $(x-c)^{k+1}$ will yield

$$f(x) = (x-c)^k q_k(x) + r_k(x)$$
$$f(x) = (x-c)^{k+1} q_{k+1}(x) + r_{k+1}(x)$$

with

$$\begin{aligned} r_k(x) &= f(c) + f'(c)(x-c) + \cdots + \frac{f^{(k-1)}(c)}{(k-1)!}(x-c)^{k-1} \\ r_{k+1} &= f(c) + f'(c)(x-c) + \cdots + \frac{f^{(k)}(c)}{k!}(x-c)^k \\ &= r_k(x) + \frac{f^{(k)}(c)}{k!}(x-c)^k \end{aligned}$$

both expressed as polynomials in $(x-c)$.

By definition $c$ is a $k$-fold root of $f(x) = 0$, if and only if $f(x)$ is divisible by $(x-c)^k$ and $f(x)$ is not divisible by $(x-c)^{k+1}$. In other words, $c$ is a $k$-fold root if and only if $r_k(x) = 0$ and $r_{k+1}(x) \neq 0$. Using the above expressions of $r_k(x)$ and $r_{k+1}(x)$, we see that $r_k(x) = 0$ and $r_{k+1}(x) \neq 0$ if and only if

$$f(c) = f'(c) = \cdots = f^{(k-1)}(c) = 0 \quad \text{and} \quad f^{(k)}(c) \neq 0 \ .$$

We have therefore proved the following characterization of multiple roots in terms of derivatives.

7.3.1 THEOREM. *Let $f(x) = 0$ be an equation with real coefficients. Then $r$ is a $k$-fold root of the equation if and only if*

$$f(r) = f'(r) = \cdots = f^{(k-1)}(r) = 0 \quad \textit{and} \quad f^{(k)}(r) \neq 0 \ .$$

In particular, $r$ is a simple root if and only if $f(r) = 0$ and $f'(r) \neq 0$. Similarly $r$ is a double root if and only if $f(x) = f'(r) = 0$ and $f''(r) \neq 0$.

Recall that the higher derivatives of $f(x)$ are defined recursively by $f^{(k)}(x) = (f^{(k-1)}(x))'$. By this remark we get the following corollary.

7.3.2 COROLLARY. *A number $r$ is a $k$-fold root of $f(x) = 0$ if and only if $f(r) = 0$ and $r$ is a $(k-1)$-fold root of $f'(x) = 0$.*

This corollary may be reformulated as follows:

7.3.3 COROLLARY. *Let $d(x) = HCF(f(x), f'(x))$. Then $r$ is a $k$-fold root of $f(x) = 0$ if and only if $r$ is a $(k-1)$-fold root of $d(x) = 0$.*

7.3.4 COROLLARY. *If $HCF(f(x), f'(x))$ is a non-zero constant, i.e. $f(x)$ and $f'(x)$ are relatively prime, then the equation $f(x) = 0$ has only simple roots.*

7.3.5 EXAMPLE. *Find the multiple roots of the equation*

$$x^3 + x^2 - 16x + 20 = 0 .$$

SOLUTIONS: Let $f(x) = x^3 + x^2 - 16x + 20$. Then $f'(x) = 3x^2 + 2x - 16$ we may use the following scheme of detached coefficients to find the HCF of $f(x)$ and $f'(x)$.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3 | 3 | 2 | $-16$ | 1 | 1 | $-16$ | 20 | $\frac{1}{3}$ |
| | 3 | $-6$ | | 1 | $\frac{2}{3}$ | $\frac{-16}{3}$ | | |
| 8 | | 8 | $-16$ | | $\frac{1}{3}$ | $\frac{-32}{3}$ | 20 | $\frac{1}{9}$ |
| | | 8 | $-16$ | | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{-16}{9}$ | |
| | | | 0 | | | $\frac{-98}{9}$ | $\frac{196}{9}$ | |

We obtain $d(x) = \text{HCF}(f(x), f'(x)) = x - 2$. Hence 2 is a simple root of $d(x) = 0$. Therefore 2 is a double root of $f(x) = 0$. Moreover a division of $f(x)$ by $(x-2)^2$ yields the quotient $x + 5$. Therefore $f(x) = 0$ has three roots, $-5$ being counted once and 2 twice.

7.3.6 REMARKS. The division algorithm for evaluating the HCF of $f(x)$ and $f'(x)$ may become very laborious as the coefficients get larger or the degree of $f(x)$ higher. Therefore the method of the example is not generally

recommended. We shall learn other more efficient methods in dealing with multiple roots.

## EXERCISE 7C

1. Show that the condition $f(r) = 0$ cannot be omitted in Corollary 7.3.2.
2. If $f(x) = (x-\alpha)^r\phi(x)$, where $r > 1$ is a positive integer and $\phi(x)$ is a polynomial such that $\phi(\alpha) \neq 0$. Show that $\alpha$ is a root of multiplicity $r-1$ of $f'(x)$.
3. By using Corollary 7.3.2, prove that the real polynomial equation $ax^2 + bx + c = 0$ has a double root if and only if $b^2 - 4ac = 0$.
4. Let $f(x) = x^3 - 3x + 2k + 8$. Find all the possible values of $k$ for which $f(x) = 0$ has repeated roots.
5. Let $f(x) = 3x^5 - 20x^3 + 45x + c$. Find all the possible values of $c$ for which $f(x) = 0$ has repeated roots.
6. If $f(x) = x^5 - 9x^4 + 26x^3 - 18x^2 + px + 27 = 0$ has an integral root of multiplicity 3. Find $p$ and solve $f(x) = 0$.
7. Given that the equation
$$x^9 + 7x^8 + 15x^7 + 9x^6 + 2x^4 + 279x^3 + 1220x^2 + 3x - 3600 = 0$$
has a repeated root, which is a negative integer, find that root.
8. Show that for real numbers $p, q$, $p \neq 0$, $x^4 + px^2 + q = 0$ has no root of multiplicity 3.
9. If the real polynomial equation $x^5 + 10a^3x^2 + b^4x + c^5 = 0$ has a real root of multiplicity 3, prove that $ab^4 - 9a^5 + c^5 = 0$.
10. Find the values of the real numbers $a$ and $b$ such that
$$(x+1)^2 | ax^4 + bx^2 + 1.$$
11. If the real polynomial equation $ax^3 + 3bx^2 + 3cx + d = 0$ has a triple root $\alpha$, show that $\frac{b}{a} = \frac{c}{b} = \frac{d}{c} = -\alpha$.
12. If real numbers $p$ and $q \neq 0$ satisfy $q^2 + 4p^5 = 0$, show that $x^5 + 5px^3 + 5p^2x + q = 0$ has 2 pairs of equal roots.

13. Show that, when $n > 3$, the equation

$$x^n + ax^2 + bx + c = 0 \qquad (c \neq 0)$$

cannot have four equal roots.

14. Given that $f_n(x) = (x+1)^n + (x-1)^n$, where $n$ is an integer greater than 1.

    (a) Show that $f_n{}'(x) = nf_{n-1}(x)$, and hence

    (b) Show that $f_n(x)$ has no multiple roots.

15. If $\alpha$ is a repeated root of $x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0 = 0$, show that $\alpha$ is a root of the equation

$$a_{n-1}x^{n-1} + 2a_{n-2}x^{n-2} + 3a_{n-3}x^{n-3} + \cdots + na_0 = 0.$$

16. Show that $x^n + nx^{n-1} + n(n-1)x^{n-2} + \cdots + n(n-1)\cdots 3\cdot 2x + n! = 0$ has no equal roots.

17. Show that the following equations have no repeated roots.

    (a) $x^n + a = 0 \qquad (a \neq 0)$

    (b) $x^6 - 6x + 1 = 0$

    (c) $x^5 + x^4 - 4x^3 + 4 = 0$

    (d) $x^6 - 4x^3 + 1 = 0$

18. Find the multiple roots of the following equations and hence solve the equations.

    (a) $x^3 - x^2 - x + 1 = 0$

    (b) $4x^4 - 4x^3 + 5x^2 - 4x + 1 = 0$

    (c) $x^5 - 5x^4 + 7x^3 + x^2 - 8x + 4 = 0$

    (d) $4x^5 + 8x^4 + x^3 - 5x^2 - x + 1 = 0$

19. If $x^n - nqx + (n-1)r = 0$ has a repeated root, where $n$ is an integer $> 1$, show that $q^n = r^{n-1}$.

20. Find real value(s) of $a$ such that $x^n + nax + n - 1 = 0$ has a repeated root, where $n$ is an integer $> 1$.

21. Let $f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$, where $n$ is a positive integer. By considering $\mathrm{HCF}(f(x), f'(x))$, show that $f(x) = 0$ has no repeated roots.

22. Let $f(x) = x^3+3px+q$, where $p$ and $q$ are real numbers. Find condition that
    (a) $f(x) = 0$ has 3 equal roots, and
    (b) $f(x) = 0$ has a pair of equal roots.
23. If the equation $x^3 + x^2 - 3bx + 3b^2 = 0$, where real number $b \neq 0$, has a multiple root, show that all the roots are identical. Hence show that if $x^4 + 4x^3 + 2x^2 - 4bx + 3b^2 = 0$ has 3 equal roots, then the fourth root is also equal.
24. Let $\alpha$ be a repeated root of $x^4 + px^3 + qx - 1 = 0$, where $p$ and $q$ are real numbers. Find $p, q$ in terms of $\alpha$ and hence, show that $(p+q)^{2/3} - (q-p)^{2/3} = (-2)^{4/3}$.
25. If $ax^3 + 3bx^2 + 3cx + d = 0$ has a repeated root, where $a, b, c$ and $d$ are real numbers such that $a \neq 0$, and $ac - b^2 \neq 0$, show that

$$(bc - ad)^2 = 4(ac - b^2)(bd - c^2).$$

26. (a) Show that $a$ is a root of

$$(a+b)x^3 + 2a(a-b)x^2 - 3a^2(a+b)x + 4a^3b = 0,$$

    where $a$ and $b$ are real numbers.

    (b) By using (a) or otherwise, show that

$$x^4 - (a+b)x^3 - a(a-b)x^2 + a^2(a+b)x - a^3b = 0$$

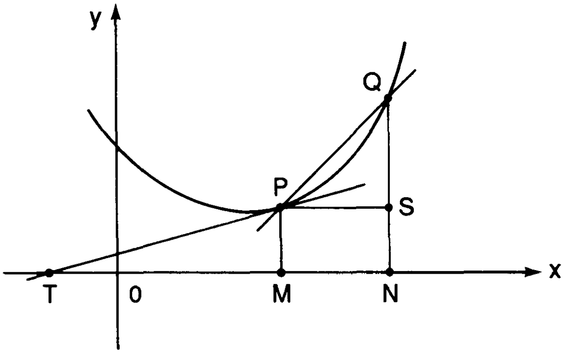
    has equal roots and hence solve the equation.
27. Let $f(x)$ and $g(x)$ be real polynomials without multiple roots and common roots. Polynomial $p(x)$ and $q(x)$ are defined by

$$p(x) = f(x)g(x), \quad F(x) = p(x)g(x)^{k+1} \quad \text{and} \quad F'(x) = q(x)g(x)^k,$$

    where $k$ is a positive integer. Show that $p(x)$ and $q(x)$ have no common roots.

## 7.4 Tangent

Given a real number $c$ and a polynomial $f(x)$. The derivative $f'(c)$ at $x = c$ of $f(x)$ is the slope of the tangent to the curve $y = f(x)$ at the point $(c, f(c))$. This geometric interpretation is based on the diagram below.




On the curve which represents $y = f(x)$, let $P$ be the point corresponding to the value $c = OM$ and $TP$ be the tangent to the curve at $P$. Take a second point $Q$ on the curve corresponding to the value $c + h = ON$ where $h$ represents a small increment. Then the lengths of the various segments in the diagram are

$$OM = c;\ MN = h;\ ON = c + h;\ MP = f(c);\ NQ = f(c + h)\ .$$

When $h$ tends towards 0, the point $Q$ approaches $P$. The chord $PQ$ will ultimately become the tangent $TP$ to the curve at $P$, and the slope of $PQ$ becomes the slope of the tangent $TP$:

$$\text{Slope of } TP = \lim_{Q \to P} \text{slope of } PQ = \lim_{h \to 0} \frac{f(c + h) - f(c)}{h}\ .$$

On the other hand, by Taylor's Formula we have

$$\frac{f(c + h) - f(c)}{h} = f'(c) + \frac{f''(c)}{2!}h + \frac{f'''(c)}{3!}h^2 + \cdots\ .$$

Therefore

$$\lim_{h \to 0} \frac{f(c + h) - f(c)}{h} = f'(c)$$

showing that $f'(c)$ is the slope of the tangent $TP$.

7.4.1 THEOREM. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$ and $f'(x)$ its derivative. The value $f'(c)$ of $f'(x)$ at $c$ is the slope of the tangent to the curve $y = f(x)$ at the point $(c, f(c))$.*

7.4.2 EXAMPLE. *Find the equation of the tangent to the curve $y = 8x^3 - 22x^2 + 13x - 2$ at the point $P$ corresponding to the $x = 1$.*

SOLUTION: Let $f(x) = 8x^3 - 22x^2 + 13x - 2$. The co-ordinates of $P$ are $(1, f(1)) = (1, -3)$. We differentiate $f(x)$ to get $f'(x) = 24x^2 - 44x + 13$ and $f'(1) = -7$. The point-slope form of the tangent at $P$ is therefore

$$y + 3 = -7(x - 1) \qquad \text{or} \qquad 7x + y + 10 = 0 \ .$$

7.4.3 EXAMPLE. *Find the tangents to the curve $y = x^3 - 3x^2 - 18x + 20$ which have slope $-9$.*

SOLUTION: $f'(x) = 3x^2 - 6x - 18$. The roots of the equation $3x^2 - 6x - 18 = -9$ are $-1$ and $3$. Therefore at $(-1, 34)$ and $(3, -34)$ the tangents to the curve have slope $-9$. The equations of these tangents are

$$9x + y - 25 = 0 \qquad \text{and} \qquad 9x + y + 7 = 0 \ .$$

7.4.4 EXAMPLE. *Find the points on the curve $y = x^3 + 3x^2 - 9x - 11$ at which the tangent is horizontal.*

SOLUTION: Let $f(x) = x^3 + 3x^2 - 9x - 11$. Then $f'(x) = 3x^2 + 6x - 9 = (3x + 9)(x - 1)$. Therefore at $P = (-3, f(-3)) = (-3, 16)$ and $Q = (1, -16)$ the curve has horizontal tangents.

## EXERCISE 7D

1. Find the equation of the tangents to the following curve at the point $P$ corresponding to $x = 1$.
   (a) $y = x^3 - 3x^2 + x - 1$.
   (b) $y = 2x^3 - x^2 + 5x - 6$.

2. Find the tangents to the curve $y = 2x^3 + 5x^2 + 8x - 10$ which have slope 12.

3. Find the points on the following curve at which the tangent is horizontal.

   (a) $y = 4x^3 + 3x^2 - 36x + 2$.

   (b) $y = x^4 + 4x^3 - 16x$.

4. For the curve $y = x^2$, find the equation of the tangent at $x = t$. Hence, find the condition for the line $\ell x + my + n = 0$ to be a tangent to the curve $y = x^2$.

5. Find the equation of the tangent to $y = x^2$ through the point $(7, 49)$. Use this tangent line to estimate $\sqrt{50}$, corrected to 3 decimal places.

6. Let $f(x)$ be a real polynomial and $y = mx + n$ is tangent to $y = f(x)$ at the point $(a, f(a))$. Show that $a$ is a double root of the equation $f(x) - mx - n = 0$.

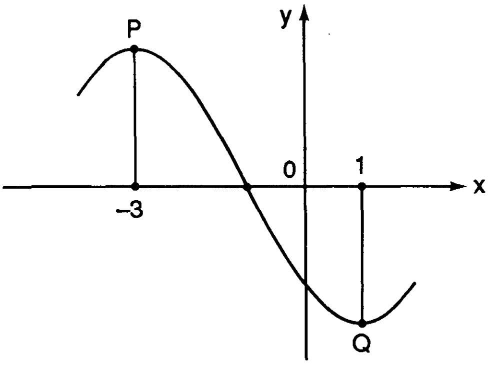
7. Let $f(x) = x^4 + ax^3 + bx^2$, where $a$ and $b$ are real numbers such that $a^2 - 4b > 0$. Show that if $y = mx + n$ is tangent to $y = f(x)$ at the points corresponding to $x = \alpha$ and $x = \beta$, $\alpha \neq \beta$, then $\alpha$, $\beta$ are the roots of

$$x^2 + \frac{a}{2}x - \frac{a^2 - 4b}{8} = 0 .$$

   Hence express the equation $y = mx + n$ in terms of $a$ and $b$.

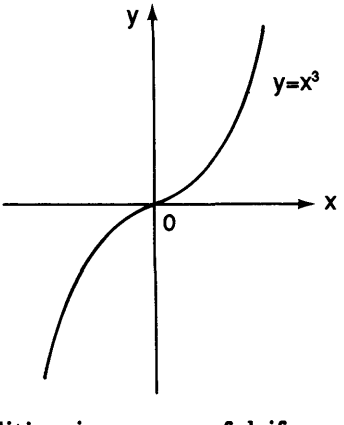
## 7.5 Maximum and minimum

The curve $y = x^3 + 3x^2 - 9x - 11$ of Example 7.4.4 is sketched below.



We see from the diagram that around the point $P = (-3, 16)$, all the adjacent points on the curve lie below the horizontal tangent and around the point $Q = (1, -16)$, all the adjacent points on the curve lie above the horizontal tangent. In the language of calculus, we say that the function $f(x)$ has a local maximum and a local minimum at $x = -3$ and $x = 1$ respectively, because $f(-3) = 16$ is the maximum value of $f(x)$ in a neighbourhood of $x = -3$ and $f(1) = -16$ is the minimum value of $f(x)$ in a neighbourhood of $x = 1$.

Let us formulate the above discussion in terms of the derivatives. We shall say that the function $f(x)$ has a *local maximum* or *local minimum* at $x = c$ if (i) the curve $y = f(x)$ has a horizontal tangent at $x = c$ and (ii) all points on the curve in the proximity of the point of contact lie on one side of the tangent. By the results of the last section, condition (i) is satisfied if and only if $f'(c) = 0$. Therefore for $f(x)$ to have a local maximum or minimum at $x = c$, it is *necessary* that $f'(c) = 0$. However the vanishing of $f'(x)$ at $x = c$ may not be sufficient for $f(x)$ to have a local maximum or minimum. To see this, we consider $f(x) = x^3$, for example. Then $f'(x) = 3x^2$. Therefore $f'(0) = 0$ and the curve $y = x^3$ has a horizontal tangent at $0 = (0, 0)$. But the adjacent points of $O$ on the curve lie on both sides of the horizontal tangent. Therefore $f(x) = x^3$ does not have a local maximum or minimum at $x = 0$ though $f'(0) = 0$.




A sufficient condition is more useful if we are interested in the locations of the local maxima and minima of a polynomial. The following theorem provides us with one such condition.

7.5.1 THEOREM. *Let $f(x)$ be a polynomial. Then $f(x)$ has a local maximum at $x = c$ if $f'(c) = 0$ and $f''(c) < 0$.*

PROOF: Suppose that $f'(c) = 0$ and $f''(c) < 0$. It is sufficient to show that $f(c) > f(c+h)$ and $f(c) > f(c-h)$ for all small increments $h > 0$. By Taylor's formula, we get

$$\begin{aligned} f(c+h) - f(c) &= \frac{f''(c)}{2!}h^2 + \frac{f'''(c)}{3!}h^3 + \frac{f''''(c)}{4!}h^4 + \cdots \\ &= h^2\Big\{\frac{f''(c)}{2!} + \frac{f'''(c)}{3!}h + \frac{f''''(c)}{4!}h^2 + \cdots\Big\}\,. \end{aligned}$$

By Corollary 6.2.2, for sufficiently small positive values of $h$, the value of the expression within the braces has the same sign as its first term $f''(c)/2!$. Since $h > 0$ and $f''(c) < 0$, it follows that $f(c) > f(c+h)$. Similarly we can show that $f(c) > f(c-h)$. Therefore $f(x)$ has a local maixmum at $x = c$.

Using the same argument we obtain a parallel result:

7.5.2 THEOREM. *Let $f(x)$ be a polynomial. Then $f(x)$ has a local minimum at $x = c$ if $f'(c) = 0$ and $f''(c) > 0$.*

Let us apply these two theorems to the curve $y = x^3+3x^2-9x+11$ which we used at the beginning of the present section. Differentiate $f(x) = x^3+3x^2-9x+11$ to get $f'(x) = 3x^2+6x-9$ and $f''(x) = 6x+6$. Therefore $f'(-3) = 0$, $f''(-3) < 0$ and $f'(1) = 0$, $f''(1) > 0$. This confirms that the curve has a local maximum at $P = (-3, 16)$ and a local minimum at $Q = (1, -16)$.

7.5.3 EXAMPLE. *Find the local maxima and minima of the polynomial $f(x) = 3x^3 - 3x^2 - 36x + 14$.*

SOLUTION: $f'(x) = 6x^2-6x-36$; $f''(x) = 12x-6$. Thus at $x = -2$ we have $f(-2) = 50$, $f'(-2) = 0$, $f''(-2) = -30$; at $x = 3$, we have $f(3) = -40$, $f'(3) = 0$, $f''(3) = 30$. Therefore $f(x)$ has a local maximum at $x = -2$ with value 50 and a local minimum at $x = 3$ with value $-40$.

7.5.4 REMARKS. The conditions of Theorems 7.5.1 and 7.5.2 are sufficient for $f(x)$ to have a local maximum and a local minimum respectively at $x = c$. Neither of them is a necessary condition. To see that, we consider $g(x) = x^4$. Here we have $g'(x) = 4x^3$, $g''(x) = 12x^2$. Therefore $g'(0) = g''(0) = 0$ and the condition of 7.5.2 is not satisfied; yet $g(x)$ has a local minimum at $x = 0$.

7.5.5 REMARKS. The local minimum of $g(x) = x^4$ at $x = 0$ is clearly a minimum of the function $g(x) = x^4$ because $g(c) \geq 0$ for all $c \in \mathbf{R}$. However for the polynomial $f(x) = x^3 + 3x^2 - 9x - 11$ studied at the beginning of this section, we have $f(-10) < f(1)$ and $f(10) > f(-3)$. Therefore $f(x)$ does not have a maximum but a local maximum at $x = -3$ nor a minimum but a local minimum at $x = 1$. Therefore we shall retain the adjective 'local' in our discussion to avoid confusion. In the literature the term *local extremum* is also used which means either a local maximum or a local minimum.

## EXERCISE 7E

1. Find the local maxima and local minima of the following polynomials.
   (a) $f(x) = x^3 - 6x^2 + 9x - 3$.
   (b) $f(x) = x^4 - 2x^2$.
   (c) $f(x) = 6x^5 - 75x^4 + 350x^3 - 750x^2 + 720x + 1$.
2. Given real polynomial $f(x) = ax^2 + bx + c$, prove that $x = -\frac{b}{2a}$ is a local minimum of $f(x)$ if $a > 0$ and a local maximum if $a < 0$.
3. Show that if $a < 0$, $f(x) = x^3 + 3ax + b$ has both a local maximum and a local minimum.

## 7.6 Bend point and inflexion point

From a geometric point of view, we can classify the points on a curve with horizontal tangents into two distinctive types. Let $P = (c, f(c))$ be a point on the curve $y = f(x)$ such that the tangent to the curve at $P$ is horizontal. We call $P$ a *bend point* of the curve if

all adjacent points of $P$ on the curve lie on one side of the tangent; otherwise $P$ is called an *inflexion point.* It follows from our discussion in the last section that $P$ is a bend point of $y = f(x)$ if and only if $f(x)$ has a local extremum at $x = c$. Therefore $P = (-3, 16)$ and $Q = (1, -16)$ are bend points of the cubic curve $y = x^3 + 3x^2 - 9x - 11$ of Example 7.4.4 whereas $0 = (0, 0)$ is an inflexion point of the cubic curve $y = x^3$.

We now proceed to characterize bend points and inflexion points by means of the higher derivatives of $f(x)$. Suppose that $y = f(x)$ has a horizontal tangent at the point $P = (c, f(c))$. Then $f'(c) = 0$; in other words, $c$ is a root of $f'(x) = 0$. Denote by $m$ $(1 \leq m \leq n-1)$ the multiplicity of this root, i.e.

$$f'(c) = f''(c) = \cdots f^{(m)}(c) = 0 \quad \text{and} \quad f^{(m+1)}(c) \neq 0 \ .$$

This special property of the derivatives simplifies Taylor's formula for $f(x)$ into

$$\begin{aligned} f(c+h) - f(c) &= h^{m+1}\Big\{\frac{f^{(m+1)}(c)}{(m+1)!} + \frac{f^{(m+2)}(c)}{(m+2)!}h + \frac{f^{(m+3)}(c)}{(m+3)!}h^2 + \cdots\Big\} \\ f(c-h) - f(c) &= (-h)^{m+1}\Big\{\frac{f^{(m+1)}(c)}{(m+1)!} - \frac{f^{(m+2)}(c)}{(m+2)!}h + \frac{f^{(m+3)}(c)}{(m+3)!}h^2 \\ &\qquad - \cdots\Big\} \ . \end{aligned}$$

The argument that we used in the proof of Theorem 7.5.1 shows that the term $A = f^{(m+1)}(c)/(m+1)!$ will be predominant for both expressions within the braces when $h > 0$ is sufficiently small. In other words for sufficiently small positive values of $h$, we may, for all practical purposes, take

$$\begin{aligned} &f(c+h) - f(c) \text{ to be } Ah^{m+1} \\ &f(c-h) - f(c) \text{ to be } (-1)^{m+1}Ah^{m+1} \end{aligned}$$

But this means that all adjacent points of $P$ on the curve will lie on *one* side of the horizontal tangent if and only if $m$ is *odd*, and the adjacent points of $P$ will be on *both* sides of the tangent if and only if $m$ is *even*. Therefore we have proved the following extension of Theorems 7.5.1 and 7.5.2.

7.6.1 THEOREM. *Let $f(x)$ be a polynomial with real coefficients. Then a point $P = (c, f(c))$ on the curve $y = f(x)$ is a bend point of the curve if and only if $c$ is a root of $f'(x) = 0$ of odd multiplicity; $P$ is an inflexion point if and only if $c$ is a root of $f'(x) = 0$ of even multiplicity.*

Applying Theorem 7.6.1 to the quartic curve $y = x^4$, we see that $O = (0,0)$ is a bend point since 0 is a triple root of $4x^3 = 0$. Therefore $f(x) = x^4$ has a local extremum at $x = 0$.

## EXERCISE 7F

1. Let $f(x)$ be a real polynomial and $P(c, f(c))$ is a bend point of the curve $y = f(x)$ such that $f'(c) = f''(c) = \cdots = f^{(m)}(c) = 0$ and $f^{(m+1)}(c) \neq 0$ for positive integer $m$. Show that $f(x)$ has a local maximum or minimum at $x = c$ according to $f^{(m+1)}(c) < 0$ or $f^{(m+1)}(c) > 0$.
2. For each of the following curves, $y = f(x)$, find the points at which the tangent is horizontal. Determine whether the points are local maximum, local minimum or point of inflection. Hence, sketch the graph and determine the number of distinct real roots for the equation $f(x) = 0$.
   (a) $y = 2x^3 - 5x^2 - 4x$.
   (b) $y = 4x^3 + 3x^2 - 36x + 2$.
   (c) $y = x^4 + 4x^3 - 16x + 2$.
   (d) $y = x^4 + 4x^3 + 6x^2 + 4x + 3$.
3. Let $f(x) = (x-a)^n g(x)$, where $f(x)$ and $g(x)$ are real polynomials, $n$ is a positive odd integer greater than 1, and $g(a) \neq 0$. Show that $(a, 0)$ is a point of inflection of the graph $y = f(x)$.
4. Determine the coefficients $a$, $b$, $c$ and $d$ such that $f(x) = ax^3 + bx^2 + cx + d$ has a local maximum at $(-1, 10)$ and an inflection point at $(1, -6)$.

CHAPTER EIGHT

# POLYNOMIALS AS CONTINUOUS FUNCTIONS

In the last chapter we treat polynomials as differentiable functions and study their derivatives and Taylor's expansions. As each differentiable function is also continuous, polynomials are continuous functions. In this chapter we shall first introduce the general concept of continuous function and prove that polynomial functions are continuous. Thus every polynomial together with all its derivatives is a continuous function. Then we shall discover some very important properties of continuous functions which are useful in the theory of equations. Readers who are not familiar with the fundamental properties of real numbers and convergence may experience difficulty in reading some proofs given in the chapter. However, if it is accepted that the graph of a polynomial is a continuous unbroken curve, which can be drawn without lifting the pencil off the paper, there should be no obstacle in the understanding of the idea of the theorems.

## 8.1 Continuity

Given a polynomial of $f(x)$ of $\mathbf{R}[x]$. To sketch the graph of the function $f(x)$, we usually proceed in the following two steps:

(1) We choose a finite series of consecutive values $c_i$ for the variable $x$ and calculate the corresponding functional values $f(c_i)$. Then we plot the points $P_i$ with coordinates $(c_i, f(c_i))$ on the Cartesian plane.

(2) We join each point $P_i$ with the next point $P_{i+1}$ by an arc to trace out a smooth continuous curve on the plane.

The curve which consists of an infinite number of points on the plane will be a rough sketch of the graph of the given polynomial function $f(x)$. Naturally the accuracy of the sketch will depend on the number of points $P_i$ that we use in the first step. However, no matter how

much work we put into this step, we can only obtain a finite number of isolated points on the graph of $f(x)$. To get a smooth curve which consists of an infinite number of points, we have to link up each $P_i$ with $P_{i+1}$ by an arc in the second step. Surely we must assume that the graph of $f(x)$ does not have any 'break' or 'jump' to justify this.

We shall see later in Section 8.3 that such an assumption is correct because the polynomial function $f(x)$ is a continuous function. As such, $f(x)$ will have a graph which is an unbroken curve. The following theorem is a precise formulation of this state of affairs given in terms of the local behaviour around each point of its graph.

8.1.1 THEOREM. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$ and $c$ be any point on the real line $\mathbf{R}$. Then for any given positive value $\varepsilon$, no matter how small, a positive value $\delta$ can be found that satisfies the following condition*

$$|f(c+h) - f(c)| < \varepsilon \quad \textit{for all} \quad |h| < \delta \ .$$

Let us read the statement of the theorem carefully before we proceed to prove it. It consists of three parts, namely:

(A) Given are the following data:
  (i) an arbitrary polynomial $f(x)$ with real coefficients
  (ii) an arbitrary real number $c$ taken from the domain of $f(x)$, and
  (iii) an arbitrary positive quantity $\varepsilon$, which is usually chosen small.

(B) To be found is another positive quantity $\delta$. This quantity $\delta$ will depend on $f(x)$, $c$ and $\varepsilon$.

(C) The quantity $\delta$ being sought shall satisfy a specific condition.

Clearly, of these three parts, only the condition in (C) requires further elaboration. To make it less concise we may expand it as follows: *The positive quantity $\delta$ should be such that for any positive number $h$ less than $\delta$, the inequalities*

$$f(c) - \varepsilon < f(c+h) < f(c) + \varepsilon$$

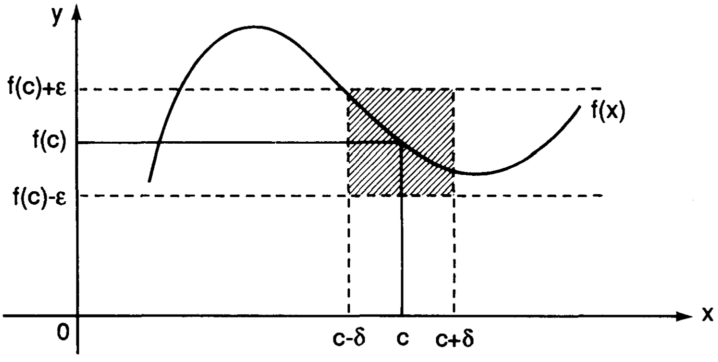
*should hold.*

Alternatively we may rephrase it into: *for all points $d$ of the domain, as long as $c - \delta < d < c + \delta$,*

$$f(c) - \varepsilon < f(d) < f(c) + \varepsilon$$

*holds for the functional values $f(c)$ and $f(d)$.*

The last version of the condition may be interpreted as follows: *As long as $d$ deviates less than $\delta$ from $c$, $f(d)$ will deviate less than $\varepsilon$ from* $f(c)$. This state of affair can be illustrated by the diagram below:




The requirement on $\delta$ is simply that the portion of the graph lying above the open interval $(c - \delta, c + \delta)$ must fall entirely within the shaded rectangle.

For example, if $f(x) = x^2 + 2x + 2$, $c = 2$, and $\varepsilon = 1$, then $|f(2+h) - f(2)| = |6h + h^2|$. In order that $|6h + h^2| < 1$, we may take $|h| < 1/10$. Thus $\delta = 1/10$ is one such positive value that will satisfy the condition of (C).

Let us now proceed to prove Theorem 8.1.1.

PROOF: Denoting $f^{(k)}(c)/k!$ by $D_k$ for $k = 1, \ldots, n$, we write down Taylor's expansion as

$$f(c+h) - f(c) = D_1 h + D_2 h^2 + \cdots + D_n h^n \ .$$

It is then required to find a positive number $\delta$ for the given positive number $\varepsilon$ such that

$$\varepsilon > |D_1 h + \cdots + D_n h^n| \quad \text{for all} \quad |h| < \delta \ .$$

For this purpose let us consider the polynomial

$$\varepsilon + D_1 h + \cdots + D_n h^n$$

in the indeterminate $h$ whose constant term $\varepsilon$ will become predominant for sufficiently small values of $|h|$. Therefore we need only apply Theorem 6.2.1 to this polynomial to obtain

$$\delta = \varepsilon/(\varepsilon + g) \quad \text{where} \quad g = \max\{|D_1|, |D_2|, \ldots, |D_n|\} .$$

Then by 6.2.1, for all $|h| < \delta$

$$|f(c+h) - f(c)| = |D_1 h + \cdots + D_n h^n| < \varepsilon .$$

This completes the proof.

In general a real-valued function $f(x)$ is said to be *continuous at* $c$ if $f(x)$ is defined at $x = c$ and the condition of Theorem 8.1.1 is satisfied for $x = c$; $f(x)$ is said to be *continuous* if it is continuous at every point of its domain. Thus by Theorem 8.1.1 all polynomial functions are continuous functions.

Continuous functions constitute a very large class of functions and they are probably the most useful and most studied functions of mathematics. In fact all functions that we encounter in secondary school mathematics are continuous; they include polynomial functions, rational functions, trigonometric functions, logarithmic functions and exponential functions.

## EXERCISE 8A

1. Let $f(x) = x^2 + 4x$. Find $\delta > 0$ such that

$$|f(1+h) - f(1)| < \frac{1}{10}$$

for all $|h| < \delta$.

2. Let $f(x)$ be a real polynomial. Prove that if $f(a) > 0$ for some real

number $a$, then there is $h > 0$ such that

$$f(x) > 0 \qquad \text{for } x \text{ in } (a - h, a + h) \ .$$

3. As an extension to Question 2, show that if $f(a) \neq 0$, then there is $h > 0$ such that

$$|f(x)| > 0 \qquad \text{for } x \text{ in } (a - h, a + h) \ .$$

## 8.2 Convergence of $f(c_n)$

In this section we prove a general property of continuous functions which we shall need in the next section. Recall that an infinite sequence $\langle a_n \rangle$ of real numbers $a_n$ $(n = 1, 2, \dots)$ is said to converge to a fixed real number $a$ if for every given $\varepsilon > 0$ there is an index $N$ such that $|a_n - a| < \varepsilon$ for all $n > N$. The similarity between the condition for convergence and the condition for continuity is evident enough for us to try to connect them in the following theorem.

8.2.1 THEOREM. *Let $f(x)$ be a function continuous at $x = c$, and let $\langle c_n \rangle$ be an infinite sequence of real numbers of the domain of $f(x)$. If the sequence $\langle c_n \rangle$ converges to $c$ then the sequence $\langle f(c_n) \rangle$ of functional values converges to $f(c)$.*

PROOF: Let $c_n \to c$. For each $n$ we write $c_n = c + h_n$. Then $h_n \to 0$. Now we proceed to prove that $f(c_n) \to f(c)$ under the hypothesis that $f(x)$ is continuous at $x = c$. By the definition of continuity, for any $\varepsilon > 0$, we have a $\delta > 0$ such that

$$|f(c + h) - f(c)| < \varepsilon \quad \text{for all} \quad |h| < \delta \ .$$

Since $h_n \to 0$, for the said $\delta > 0$ above there is an index $N$ such that

$$|h_n| < \delta \quad \text{for all} \quad n > N \ .$$

Therefore for the same $\varepsilon$ and $N$

$$|f(c_n) - f(c)| = |f(c + h_n) - f(c)| < \varepsilon \quad \text{for all} \quad n > N$$

since in this case $|h_n| < \delta$. Hence $f(c_n) \to f(c)$ and the proof is complete.

To put the conclusion of the theorem into a more concise form, we may say that for any convergent sequence $\langle c_n \rangle$,

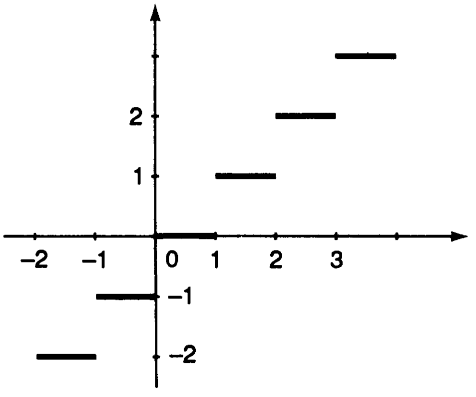
$$\lim_{n\to\infty} f(c_n) = f(\lim_{n\to\infty} c_n) \ .$$

In other words we may interchange the order of the action of taking limit and the action of calculating functional value.

Moreover we want to remark that the converse of Theorem 8.2.1 also happens to be valid. Thus a definition of continuity can be given in terms of convergence: *$f(x)$ is continuous at $x = c$ if and only if $f(c_n) \to f(c)$ as long as $c_n \to c$.* However for our purpose it is enough to know that for polynomials $f(x)$ with real coefficients, if $c_n \to c$ then $f(c_n) \to f(c)$. The converse of 8.2.1 will not be used in the sequel and we shall not give a proof thereof in order to remain on the main track of our study.

To conclude this section, we consider two functions which are not continuous.

8.2.2 EXAMPLE. For every real number $x$, we denote by $[x]$ the *integral part* of $x$, i.e. the greatest integer less than or equal to $x$. Thus $[n] = n$ for all integers $n$, $[1\frac{1}{2}] = 1$, $[-\frac{1}{2}] = -1$, $[\pi] = 3$, $[e] = 2$, etc. Define $f(x) = [x]$ for all real numbers $x$. Then $f(x)$ is a function of the set $\mathbf{R}$ into $\mathbf{R}$. Clearly if $c$ is not an integer then $f(x)$ is continuous at $x = c$. If $n$ is an integer then $f(n) = n$, $f(n + h) = n$ and $f(n - h) = n - 1$ for all $0 < h < 1$. Therefore $f(x)$ is not continuous at $x = n$. The graph $f(x)$ is given in the diagram below: $f(x)$ jumps at every integral value of $x$. Because of the shape of its graph, $f(x)$ is called a *step function.*




8.2.3 EXAMPLE. *Consider $g(x) = \sin(1/x)$. This function is defined for every real value except at $x = 0$. If we extend the domain by assigning a real value, say $a$, to $g(0)$, then the function $g(x)$ would be defined at all points of $\mathbf{R}$. But it will be discontinuous at $x = 0$, whatever the value of $a$. Because if $\langle c_i \rangle$ is a null sequence, then the sequence $\langle g(c_i) \rangle$ diverges and does not converge to $f(0) = a$.*

EXERCISE 8B

Show that the following functions are not continuous at the specific points.

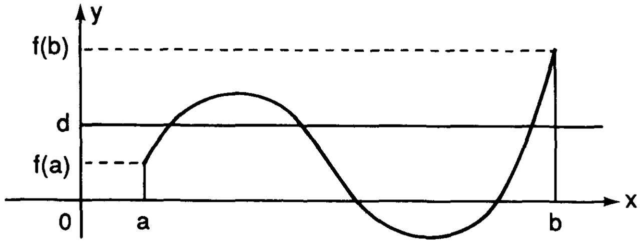
1. $f(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$ , at $x = 0$.
2. $f(x) = \begin{cases} \sin^2 \frac{1}{x} \\ 0 \end{cases}$ , at $x = 0$.
3. $f(x) = x - [x]$ at $x = n$, for all $n$ in $\mathbf{N}$.

## 8.3 Bolzano's theorem

Being a continuous function, a polynomial $f(x)$ will have many interesting and useful properties. The study of such properties constitutes an important component of the branch of mathematics called mathematical analysis. For the study of the distribution of the roots of a polynomial equation we shall need the following property of continuous functions discovered by Bernhard Bolzano (1781–1848).

8.3.1 INTERMEDIATE VALUE THEOREM. *Let $f(x)$ be a continuous function and $a < b$ two arbitrary real numbers. If $d$ lies between $f(a)$ and $f(b)$, then there is an intermediate value $c$ between $a$ and $b$ such that $f(c) = d$.*

In geometric terms, the theorem states that the portion of the curve $y = f(x)$ between the points $(a, f(a))$ and $(b, f(b))$ must cross the horizontal line $y = d$. In other words, the set $\{f(c) : a \leq c \leq b\}$ of functional values covers the entire interval between $f(a)$ and $f(b)$. Thus the theorem also means intuitively that the graph of $f(x)$ must be an unbroken curve.




PROOF: Clearly we can dispense with the trivial case in which $f(a) = f(b)$, $d = f(a)$ or $d = f(b)$. We may therefore assume that $f(a) < d < f(b)$. Redesignating $a$ by $a_1$ and $b$ by $b_1$, we consider the interval $J_1 = [a_1, b_1]$. Then by the above assumption, $J_1$ is an interval of the form $H$ such that

$$H = [s, t] \quad \text{and} \quad f(s) < d < f(t) .$$

We divide the interval into two halves $[a_1, m_1]$ and $[m_1, b_1]$ where $m_1 = \frac{1}{2}(a_1 + b_1)$. If $f(m_1) = d$, then the theorem is proved. Otherwise, either $d < f(m_1)$ or $f(m_1) < d$; hence either $f(a_1) < d < f(m_1)$ or $f(m_1) < d < f(b_1)$. Therefore either $[a_1, m_1]$ or $[m_1, b_1]$ has the form $H$. Denote that half interval by $J_2 = [a_2, b_2]$ and proceed to divide it into two quarter intervals $[a_2, m_2]$ and $[m_2, b_2]$ where $m_2 = \frac{1}{2}(a_2+b_2)$. If $f(m_2) = d$, then the theorem is proved. Otherwise one of the quarter intervals say $J_3 = [a_3, b_3]$, must have the form $H$. Further subdivisions will lead to either one of the two possible outcomes (A) or (B).

(A) We arrive at an interval $J_k = [a_k, b_k]$ which is $2^{1-k}$ of the original length and has a midpoint $m_k$ such that $f(m_k) = d$. In this case we can put $c = m_k$, and the theorem is proved.

(B) We have an infinite sequence of nested intervals

$$J_1 \supset J_2 \supset \cdots \supset J_k \supset \cdots$$

where each $J_k = [a_k, b_k]$ has the said form $H$ and is a half interval of preceding $J_{k-1} = [a_{k-1}, b_{k-1}]$. Thus the length of $J_k$ tends towards 0 as $k$ tends towards infinity. By the postulate of continuity of $\mathbf{R}$, there is a real number $c$ such that

$$a_1 \leq a_2 \leq \cdots \leq a_k \leq \cdots c \cdots \leq b_k \leq \cdots \leq b_2 \leq b_1$$

and $\lim a_k = c = \lim b_k$. Therefore

$$f(\lim a_k) = f(c) = f(\lim b_k) \ .$$

On the other hand, since all $J_k$ have the said form $H$, $f(a_k) < d < f(b_k)$. Therefore

$$\lim f(a_k) \leq d \leq \lim f(b_k) \ .$$

Now by Theorem 8.2.1, we must conclude that

$$f(c) \leq d \leq f(c)$$

since $f(\lim a_k) = \lim f(a_k)$ and $f(\lim b_k) = \lim f(b_k)$. Therefore $f(c) = d$ also in case (B). Our proof is complete.

Let us write the general Intermediate Value Theorem 8.3.1 into a form which is readily applicable to the theory of equations. For lack of a better description and easy reference we shall refer to it as Bolzano's theorem for polynomials.

8.3.2 BOLZANO'S THEOREM FOR POLYNOMIALS. *Let $f(x)$ be a polynomial of $\mathbf{R}[x]$. If for $a < b$, $f(a)$ and $f(b)$ have opposite signs, then there are an odd number of roots of $f(x) = 0$ between $a$ and $b$, each $k$-fold root being counted as $k$ roots. If $f(a)$ and $f(b)$ have the same sign, then between $a$ and $b$, $f(x) = 0$ either has no root or an even number of roots when each $k$-fold root is counted as $k$ roots.*

PROOF: Consider the case where $f(a)$ and $f(b)$ have opposite signs. Then 0 must lie between $f(a)$ and $f(b)$. By Theorem 8.3.1 $f(x) = 0$ has at least

one root between $a$ and $b$. Suppose that there are an even number of roots of $f(x) = 0$ between $a$ and $b$. Denoting these roots by $r_1, \ldots, r_{2m}$, we can write

$$f(x) = (x - r_1)(x - r_2) \cdots (x - r_{2m})g(x)$$

where $g(x) = 0$ has no root between $a$ and $b$. Now both $(a - r_1)(a - r_2) \cdots (a - r_{2m})$ and $(b - r_1)(b - r_2) \cdots (b - r_{2m})$ are positive. Hence $f(a)$ and $g(a)$ have the same sign; $f(b)$ and $g(b)$ have the same sign. Therefore $g(a)$ and $g(b)$ have opposite signs, and by 8.3.1 $g(x) = 0$ has at least one root between $a$ and $b$. But this is absurd. Therefore $f(x) = 0$ can only have an odd number of roots between $a$ and $b$. This completes the proof of the first statement. Using similar argument we can prove the second statement of the theorem.

The obvious way to apply Bolzano's theorem is to set up a table of the form

| $x$ | $c_1$ | $c_2$ | $\cdots$ | $c_{r-1}$ | $c_r$ |
|---|---|---|---|---|---|
| $f(x)$ | | | | | |

where the top row is a strictly increasing sequence of real numbers and the bottom row is filled by $+$ or $-$ signs or 0's according to whether $f(c_i)$ is positive, negative or 0. Such a table will provide us with some rough idea of the distribution of real roots of the equation $f(x) = 0$. Namely there will be at least one root of $f(x) = 0$ between $c_i$ and $c_{i+1}$ if $f(c_i)$ and $f(c_{i+1})$ have opposite signs. But it tells us nothing about the existence of roots in the intervals $(c_j, c_{j+1})$ where $f(c_j)$ and $f(c_{j+1})$ have the same sign. Because of the inherent limitation of the method we shall only have imprecise and incomplete information on the distribution of real roots of the equation. As we shall study Sturm's method in the next chapter which gives more precise information on the distribution of roots, we shall only state two easy corollaries from which quick information can be obtained.

8.3.3 COROLLARY. *If $a_n > 0$ and $n$ is odd, then the equation $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ has a root which has the opposite sign of $a_0$.*

PROOF: For convenience we denote by $f(\infty)$ and $f(-\infty)$ the value of $f(a)$ and $f(-a)$ for a sufficiently large positive value $a$ such that the leading term of $f(x)$ becomes predominant. Then the table for the equation has the form

| $x$ | $-\infty$ | $0$ | $\infty$ |
|---|---|---|---|
| $f(x)$ | $-$ | $+\ (a_0 > 0)$<br>$-\ (a_0 < 0)$ | $+$ |

Therefore by 8.3.2, if $a_0$ is negative, then $f(x) = 0$ has a positive root since $f(0)$ and $f(\infty)$ have opposite signs. Similarly $f(x) = 0$ has a negative root if $a_0$ is positive.

8.3.4 COROLLARY. *An equation $f(x) = 0$ of even degree has a positive and a negative root if the leading coefficient and the constant term of $f(x)$ have opposite signs.*

PROOF: The table for the equation has the form

| $x$ | $-\infty$ | $0$ | $\infty$ |
|---|---|---|---|
| $f(x)$ | $-$<br>$+$ | $+\ (a_0 > 0)$<br>$-\ (a_0 < 0)$ | $-$<br>$+$ |

Therefore the corollary holds.

8.3.5 EXAMPLES. *The equation $x^3 + ax^2 + bx - 3 = 0$ has a positive root. The equation $x^4 + ax^3 + bx^2 + cx - 1 = 0$ has a positive and a negative root.*

## EXERCISE 8C

1. Let $f(x) = ax^2 + bx + c$, $a \cdot c \neq 0$. Without using Corollary 8.3.4, prove that $f(x) = 0$ has a positive and a negative root if and only if $a \cdot c < 0$.
2. Let $f(x)$ be a real polynomial. Prove that for any real numbers $\alpha < \beta$, we can find an $\gamma$ in $[\alpha, \beta]$ such that $f(\gamma) = \frac{1}{2}(f(\alpha) + f(\beta))$.

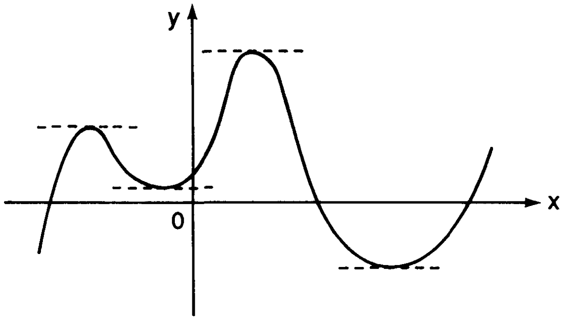
3. Let $f(x)$ be a real polynomial such that $f(x) > 0$ for all real $x$. Prove that for any real numbers $\alpha < \beta$, we can find an $\gamma$ in $[\alpha, \beta]$ such that $f^2(\gamma) = f(\alpha) \cdot f(\beta)$.
4. Given real numbers $a_1 < a_2 < \cdots < a_{2n-1} < a_{2n}$ and $k$ is real, show that $(x-a_1)(x-a_3)(x-a_5)\cdots(x-a_{2n-1}) + k^2(x-a_2)(x-a_4)\cdots(x-a_{2n}) = 0$ has $n$ distinct real roots.
5. For real numbers $a_1, a_2, \ldots, a_n$, and $b_1, b_2, \ldots, b_n$, let $f(x) = (a_1x-1)(a_2x-1)\cdots(a_nx-1)+(b_1x-1)(b_2x-1)\cdots(b_nx-1)$, where $a_1 > b_1 > a_2 > b_2 > \cdots > a_n > b_n > 0$. Show that $f(x) = 0$ has $n$ distinct real roots.
6. Given that $a$ is real, show that $(x-a)(x-(a+2))(x-(a+4))\cdots(x-(a+2n)) - 1 = 0$ has $n+1$ distinct real roots.
7. Given that $\lambda_r$, $a_r$ are real for $r = 1, 2, 3, \ldots, n$, and $a_1 < a_2 < \cdots < a_n$, show that $\lambda_1^2(x-a_2)(x-a_3)\cdots(x-a_n) + \lambda_2^2(x-a_1)(x-a_3)\cdots(x-a_n) + \cdots + \lambda_n^2(x-a_1)(x-a_2)\cdots(x-a_{n-1}) - (x-a_1)(x-a_2)\cdots(x-a_n)$ has $n$ real roots.
8. Let $f(x)$ be a real polynomial of degree greater than 2 and $\alpha$, $\beta$ be two consecutive roots of $f(x) = 0$. Show that there are an odd number of roots of $f(x) + f'(x) = 0$ in the interval $(\alpha, \beta)$, where a $k$-fold root is counted as $k$ roots.
9. Given that $p$, $q$ are distinct real roots of $(x-b)(x-c) - f^2 = 0$, where $b$, $c$, $f$ are real, $p > q$ and $f > 0$.
   (a) Show that $p > b$ and $c > q$.
   (b) If $\phi(x) = (x-a)(x-b)(x-c) - f^2(x-a) - g^2(x-b) - h^2(x-c) + 2fgh$, by considering the values of $\phi(p)$ and $\phi(q)$, show that $\phi(x) = 0$ has three real roots, where $a$, $g$, $h$ are real.

### 8.4 Rolle's theorem

As a first step towards obtaining more precise information on the distribution of real roots, we prove a theorem discovered by Michel Rolle (1652–1719) on a relationship between the roots of an equation $f(x) = 0$ and the roots of its derived equation $f'(x) = 0$.

8.4.1 ROLLE'S THEOREM. *Let $f(x)$ be a polynomial. Between two consecutive roots of $f(x) = 0$ there are an odd number of roots of $f'(x) = 0$, each $k$-fold root being counted as $k$ roots.*

Now $f(x)$ being a continuous function, we can very well imagine that if $a < b$ are two consecutive roots, then the value of $f(x)$ varying from $f(a) = 0$ to $f(b) = 0$ must begin either by increasing and then diminishing or the other way round. Therefore intuitively the curve $y = f(x)$ must have at least one bend point in between. In the proof we shall use the fact that both $f(x)$ and $f'(x)$, being polynomials, are continuous functions.




PROOF: Let $a < b$ be two consecutive roots of $f(x) = 0$. Then $f(c) \neq 0$ for all $c$ such that $a < c < b$. Taking out all factors of the form $(x - a)$ and $(x - b)$, we write

$$f(x) = (x-a)^r (x-b)^s q(x)$$

where the quotient $q(x)$ has no more root between $a$ and $b$. By Bolzano's theorem, $q(a)$ and $q(b)$ are non-zero and have the same sign. Taking derivatives we obtain

$$\frac{f'(x)}{(x-a)^{r-1}(x-b)^{s-1}} = r(x-b)q(x) + s(x-a)q(x) + (x-a)(x-b)q'(x) \ .$$

Now it follows from the fact that $q(a)$ and $q(b)$ have the same sign that the polynomial

$$h(x) = r(x-b)q(x) + s(x-a)q(x) + (x-a)(x-b)q'(x)$$

have opposite signs at $a$ and $b$. Therefore $h(x) = 0$ has an odd number of

roots between $a$ and $b$. But then it follows from $(c-a)^{r-1}(c-b)^{s-1} \neq 0$ for all $c$ in the interval $(a,b)$, and $f'(x) = (x-a)^{r-1}(x-b)^{s-1}h(x)$ that the two equations $h(x) = 0$ and $f'(x) = 0$ have the same roots in the interval $(a,b)$. Therefore $f'(x) = 0$ has an odd number of roots between two consecutive roots $a$ and $b$ of $f(x) = 0$.

For our purpose of seeking information on the distribution of roots the following corollaries are more useful than Rolle's theorem.

8.4.2 COROLLARY. *Between two consecutive roots of $f'(x) = 0$ lies at most one root of $f(x) = 0$.*

PROOF: Let $r < s$ be two consecutive roots of $f'(x) = 0$. Suppose that $f(x) = 0$ has two distinct roots $a$ and $b$ so that $r < a < b < s$. Then by Rolle's theorem $f'(x) = 0$ has a root $t$ such that $a < t < b$ which is impossible, since $f'(x) = 0$ should have no root between $r$ and $s$.

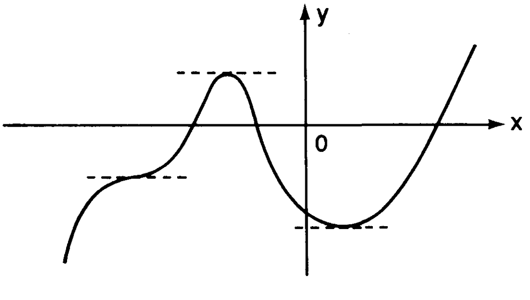
8.4.3 COROLLARY. *Let $d_1 < d_2 < \cdots < d_m$ be all the real roots of $f'(x) = 0$. Then the following statements on the distribution of roots of $f(x) = 0$ hold.*

*(i) There is at most one real root of $f(x) = 0$ greater than $d_m$ and there is at most one real root of $f(x) = 0$ less than $d_1$.*

*(ii) If $f(d_i)$ and $f(d_{i+1})$ have opposite signs, then there is exactly one real root of $f(x) = 0$ between $d_i$ and $d_{i+1}$.*

*(iii) If $f(d_i)$ and $f(d_{i+1})$ have the same sign, then there is no real root of $f(x) = 0$ between $d_i$ and $d_{i+1}$.*

*Moreover all such roots mentioned above are simple roots of $f(x) = 0$ if exist.*



PROOF: Let us first prove the concluding statement on the simplicity of any possible root of $f(x) = 0$ in the intervals between $-\infty, d_1, d_2, \ldots, d_m, \infty$. We observe that each multiple root of $f(x) = 0$ is a root of $f'(x) = 0$. Therefore the roots in the intervals must be simple.

(i) Suppose there are two roots greater than $d_m$, then by Theorem 8.4.1 there is a root of $f'(x) = 0$ between them; hence there is a root of $f'(x) = 0$ greater than $d_m$ which is impossible. Therefore $f(x) = 0$ has at most one root greater than $d_m$. Similarly $f(x) = 0$ has at most one root less than $d_1$.

(ii) If $f(d_i)$ and $f(d_{i+1})$ have opposite signs, then $f(x) = 0$ has at least one root between $d_i$ and $d_{i+1}$ by Bolzano's theorem. On the other hand if $f(x) = 0$ were to have more than one root between $d_i$ and $d_{i+1}$, then $f'(x) = 0$ would have at least one root between $d_i$ and $d_{i+1}$ which is impossible. Therefore between $d_i$ and $d_{i+1}$ there is exactly one root of $f(x) = 0$.

(iii) If $f(d_i)$ and $f(d_{i+1})$ have the same sign, then by Theorem 8.3.2 $f(x) = 0$ has either no root or an even number of roots between $d_i$ and $d_{i+1}$. In the former case (iii) holds. The latter case is impossible since $f'(x) = 0$ would have a root between $d_i$ and $d_{i+1}$.

The proof is complete.

8.4.4 EXAMPLE. *Find the intervals on the real line* $\mathbf{R}$ *in which lie the roots of the equation* $3x^5 - 25x^3 + 60x - 20 = 0$.

SOLUTION: Let $f(x) = 3x^5 - 25x^3 + 60x - 20$. Then $\frac{1}{15}f'(x) = x^4 - 5x^2 + 4 = (x^2-1)(x^2-4)$. Hence the roots of $f'(x) = 0$ are $\pm 1$ and $\pm 2$. The signs of $f(-\infty), f(-2), \ldots$ are tabulated as follows:

| $x$ | $-\infty$ | $-2$ | $-1$ | $1$ | $2$ | $\infty$ |
|---|---|---|---|---|---|---|
| $f(x)$ | $-$ | $-$ | $-$ | $+$ | $-$ | $+$ |

Thus there is one simple root in each interval $(-1, 1)$, $(1, 2)$ and $(2, \infty)$. Since $f(0) < 0$, we may replace $(-1, 1)$ by $(0, 1)$. Using Theorem 6.3.1, we may replace $(2, \infty)$ by $(2, 4)$. Therefore we conclude that the equation has two imaginary roots and one real root in each interval $(0, 1)$, $(1, 2)$ and $(2, 4)$.

8.4.5 REMARKS. The above example shows that unless the roots of the derived equation $f'(x) = 0$ are readily found, Rolle's theorem and its corollaries do not provide us with an easy means to isolate the roots of $f(x) = 0$.

## EXERCISE 8D

1. Show that $6x^5 + 15x^4 - 50x^3 - 60x^2 + 180x + 500 = 0$ has a real root in the interval $(-3, -\sqrt{2})$.
2. Find out the number of real roots of $3x^5 - 50x^3 + 135x + 20 = 0$.
3. Let $f(x) = (x-1)(x-2)(x-3)(x-4)$. By using Rolle's theorem, show that $f'(x) = 0$ has exactly three real roots and find the intervals in which the roots lie.
4. If the real polynomial $f(x)$ of degree 11 has exactly seven real roots, what are the possibilities for the number of real roots of $f'(x) = 0$?
5. Show that $4ax^3 + 3bx^2 + 2cx = a + b + c$, for real numbers $a$, $b$, and $c$, has at least one real root in $(0, 1)$.
6. Prove that if real numbers $a_0, a_1, \ldots, a_n$ satisfy $a_0 + \frac{a_1}{2} + \cdots + \frac{a_n}{n+1} = 0$, then $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ has a real root in $(0, 1)$.
7. If $a_4x^4 + a_3x^3 + a_2x^2 + a_1x = 0$ has a positive real root $x = \alpha$, prove that the equation $4a_4x^3 + 3a_3x^2 + 2a_2x + a_1 = 0$ has a positive real root smaller than $\alpha$.
8. Prove that the equation $x^3 - 3x + c = 0$ never has two real roots in $[0, 1]$, no matter what real value of $c$ may be.
9. The equation in Question 3 may have real roots elsewhere. For $f(x) = x^3 - bx + c$, $b > 0$ and $4b^3 - 27c^2 > 0$, show that $f(x) = 0$ has 3 real roots.
10. Prove that $x^3 - x^2 + 2x + c = 0$ has only one real root no matter what $c$ may be.
11. Let $f(x) = (x^2 - 1)^4$.
    (a) Show that $f'(x) = 0$ has seven real roots.
    (b) Show that the roots of $f^{(4)}(x) = 0$ are real and distinct.
12. Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ in $\mathbf{R}[x]$. If $b_1 < b_2 < \cdots < b_k$ are the distinct real roots of $f(x) = 0$ with multiplicities

$m_1, m_2, \ldots, m_k$ respectively, then show that $f'(x) = 0$ has at least $(m_1 + m_2 + \cdots + m_k) - 1$ roots in $(b_1, b_k)$.

13. Let $a_0, a_1, a_2, \ldots, a_5$ be real numbers such that the equation $a_5x^5 + \cdots + a_2x^2 + a_1x + a_0 = 0$ has five real roots. Show that the roots of $36a_5x^5 + 25a_4x^4 + 16a_3x^3 + 9a_2x^2 + 4a_1x + a_0 = 0$ are all real.

14. Prove that the $n$-th Legendre polynomial
$$p_n(x) = \frac{1}{2^n \cdot n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$
has $n$ simple real roots, each lying in $(-1, 1)$.

15. Let $f(x) = x^{2n} - 2nx + (2n - 1)r$, where $r$ is a real number and $n$ is a positive integer. Find the range of $r$ for which $f(x) = 0$ has real roots.

16. With a slight modification of $f(x)$ in Question 15, suppose $f(x) = x^{2n+1} - (2n + 1)x + 2nr$, find the range of $r$ for which $f(x) = 0$ has three real roots.

17. Consider $f(x) = x^3 + a_2x^2 + a_1x + a_0$ of $\mathbf{R}[x]$. If for some real numbers $a < b$, $f'(a) = f'(b) = 0$, determine the number of real roots of $f(x) = 0$ if

    (a) $f(a)f(b) > 0$,

    (b) $f(a)f(b) = 0$, or

    (c) $f(a)f(b) < 0$.

18. For real numbers $a < b$ and a real polynomial $f(x)$, prove that if $f(b) = 0$, $f(a)f'(a) > 0$ and $f(x) \neq 0$ in $(a, b)$, then $f'(x) = 0$ has at least one real root in $(a, b)$.

19. Let $f(x) = a_nx^n + \cdots + a_{s+1}x^{s+1} + a_sx^s + a_0$, where $n > s$ are positive integers with $a_n \neq 0$ and $a_0a_s > 0$. If $b$ is the smallest positive root of $f(x) = 0$, show that $f'(x) = 0$ has a root in $(0, b)$.

20. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ in $\mathbf{R}[x]$. If all $a_i$'s are non-zero, $c_1$ and $c_2$ are the number of changes of sign of coefficients of $f'(x)$ and $f(x)$ respectively, $p_1$ and $p_2$ are the number of positive roots of $f'(x) = 0$ and $f(x) = 0$ respectively, show that if $c_1 \geq p_1$, then $c_2 \geq p_2$.

21. Let $f(x)$ be a real polynomial of degree $n$ with simple roots only, $d_1 < d_2 < \cdots < d_m$ be all the real roots of $f'(x) = 0$ and $f(d_i) \neq 0$ for $i = 1, 2, \ldots, m$ $(< n)$.

(a) Show that $f(x) = 0$ has at most $m + 1$ real roots.

(b) Show that, for $k < n$, if $f^{(k)}(x)$ has imaginary roots, then $f(x) = 0$ also has imaginary roots.

22. Let $f_k(x) = 1 + x + \frac{x^2}{2!} + \frac{x^2}{3!} + \cdots + \frac{x^k}{k!}$, $k$ in $\mathbf{N}$.

(a) Show that $f_k'(x) = f_{k-1}(x)$.

(b) Show that if $f_{2k-1}(\alpha) = 0$, then $f_{2k}(\alpha) > 0$.

(c) Show, by mathematical induction, that $f_{2k}(x) = 0$ has no real root and $f_{2k-1}(x)$ has one negative real root for any positive integer $k$.

CHAPTER NINE

# SEPARATION OF REAL ROOTS

The method of separation of roots based on Rolle's theorem of the last chapter has one major disadvantage in that the roots of the equation $f'(x) = 0$ have to be found before the roots of $f(x) = 0$ can be isolated. Now if $\deg f(x) = n$, then $\deg f'(x) = n - 1$. For large $n$, it is far more difficult to find the exact values of the real roots of the equation $f'(x) = 0$ than to separate the roots of $f(x)$ by intervals. In this chapter we shall study three useful methods of separation, the best of which is discovered by the Swiss mathematician Jacques Sturm (1803–1855).

## 9.1 The Sturm sequence

Recall that given two positive integers $a$ and $b$ their greatest common divisor $\gcd(a, b)$ can be evaluated by a standard alternate division algorithm:

$$\begin{aligned} a &= bq_1 + r_1 & \quad 0 \le r_1 < b \\ b &= r_1 q_2 + r_2 & \quad 0 \le r_2 < r_1 \\ & \cdots\cdots\cdots \\ r_{m-1} &= r_m q_{m+1} \\ r_m &= \gcd(a, b)\ . \end{aligned}$$

Similarly given two non-zero polynomials $f(x)$ and $g(x)$, we can carry out a series of successive Euclidean algorithms:

$$\begin{aligned} f(x) &= g(x)q_1(x) + r_1(x), \quad r_1(x) = 0 \text{ or } \deg r_1(x) < \deg g(x) \\ g(x) &= r_1(x)q_2(x) + r_2(x), \quad r_2(x) = 0 \text{ or } \deg r_2(x) < \deg r_1(x) \\ & \cdots\cdots\cdots \end{aligned}$$

$$r_{k-1}(x) = r_k(x)q_{k+1}(x) + r_{k+1}(x), \quad r_{k+1} = 0 \text{ or } \deg r_{k+1}(x) < \deg r_k(x)$$

$$\cdots\cdots\cdots$$

Because the successive remainders $r_k(x)$ have strictly decreasing degrees, the process has to terminate after a finite number of steps at which we obtain

$$r_{m-1}(x) = r_m(x)q_{m+1}(x)$$

with a vanishing remainder $r_{m+1}(x) = 0$. On the other hand it follows from Theorem 2.3.2(e) that for each step of the process we have

$$\mathrm{HCF}(r_{k-1}(x), r_k(x)) = \mathrm{HCF}(r_k(x), r_{k+1}(x)) \ .$$

Telescoping these equations, we obtain

$$\begin{aligned} &\mathrm{HCF}(f(x), g(x)) \\ &= \mathrm{HCF}(g(x), r_1(x)) = \mathrm{HCF}(r_1(x), r_2(x)) = \cdots \\ &= \mathrm{HCF}(r_{m-2}(x), r_{m-1}(x)) = \mathrm{HCF}(r_{m-1}(x), r_m(x)) = r_m(x) \ . \end{aligned}$$

Let us apply this process to a given polynomial $f(x)$ and its first derivative $f'(x)$. The first step of the algorithm gives

$$f(x) = f'(x)q(x) + r(x)$$

where the remainder $r(x)$ is either the zero polynomial or has a degree strictly less than that of $f'(x)$. According to the scheme proposed by Sturm, we modify each step of the algorithm by a change of sign of the remainder:

$$f(x) = f'(x)q(x) - (-r(x)) \ .$$

Also instead of $r(x)$, the negative $-r(x)$ of the remainder will be used as the divisor in the next step. The result of each division will be written as dividend equals divisor times quotient minus the negative of the remainder. Thus denoting $f(x)$ by $f_0(x)$, $f'(x)$ by $f_1(x)$, $-r(x)$ by $f_2(x)$, we write the first two steps as follows:

$$\begin{aligned} f_0(x) &= f_1(x)q_1(x) - f_2(x) \\ f_1(x) &= f_2(x)q_2(x) - f_3(x) \end{aligned}$$

where $-f_3(x)$ is just the remainder of second division. Similarly the next step will give

$$f_2(x) = f_3(x)q_3(x) - f_4(x)$$

and in general

$$f_{k-1}(x) = f_k(x)q_k(x) - f_{k+1}(x) .$$

We carry out this slightly modified process of alternate divisions until a vanishing $-f_{m+1}(x) = 0$ is obtained. Thus we have, at the end, a sequence of non-zero polynomials

$$f_0(x), f_1(x), f_2(x), \dots , f_m(x)$$

of strictly decreasing degrees where

$$f_0(x) = f(x), \quad f_1(x) = f'(x) \quad \text{and} \quad f_m(x) = \mathrm{HCF}(f(x), f'(x)) .$$

This sequence is called a *Sturm sequence* or a sequence of *Sturm functions* of $f(x)$.

9.1.1 EXAMPLE. *For the polynomial $f(x) = x^3 + 4x^2 - 8$ the above modified process applied on $f_0(x) = f(x) = x^3 + 4x^2 - 8$ and $f_1(x) = f'(x) = 3x^2 + 8x$ will give*

$$\begin{aligned} x^3 + 4x^2 - 8 &= (3x^2 + 8x)(\frac{1}{3}x + \frac{4}{9}) - (\frac{32}{9}x + 8) \\ 3x^2 + 8x &= (\frac{32}{9}x + 8)(\frac{27}{32}x + \frac{45}{128}) - (\frac{45}{16}) . \end{aligned}$$

*Thus we have a Sturm sequence as follows:*

$$\begin{aligned} f_0(x) &= x^3 + 4x^2 - 8 \\ f_1(x) &= 3x^2 + 8x \\ f_2(x) &= \frac{32}{9}x + 8 \\ f_3(x) &= \frac{45}{16} . \end{aligned}$$

*Furthermore since HCF$(f(x), f'(x)) = f_3(x)$ is a non-zero constant, by Corollary 7.3.4, the equation $f(x) = 0$ has no multiple root.*

9.1.2 EXAMPLE. *Find a sequence of Sturm functions of the polynomial* $f(x) = x^4 - 6x^3 + 13x^2 - 12x + 4$.

SOLUTION: $f'(x) = 4x^3 - 18x^2 + 26x - 12$. Then

$$f(x) = f'(x)(\frac{1}{4}x - \frac{3}{8}) - (\frac{1}{4}x^2 - \frac{3}{4}x + \frac{1}{2})$$
$$f'(x) = (\frac{1}{4}x^2 - \frac{3}{4}x + \frac{1}{2})(16x - 24) .$$

Therefore we have a sequence of Sturm functions as follows,

$$f_0(x) = x^4 - 6x^3 + 13x^2 - 12x + 4$$
$$f_1(x) = 4x^3 - 18x^2 + 26x - 12$$
$$f_2(x) = \frac{1}{4}(x^2 - 3x + 2) .$$

Since $\mathrm{HCF}(f(x), f'(x)) = f_2(x) = \frac{1}{4}(x^2 - 3x + 2) = \frac{1}{4}(x-2)(x-1)$, we see that both 1 and 2 are double roots of the equation $f(x) = 0$. Hence $x^4 - 6x^3 + 13x^2 - 12x + 4 = (x-1)^2(x-2)^2$.

## EXERCISE 9A

1. Find a Sturm sequence for $f(x) = x^4 - 5x^2 + 8x - 8$, and hence show that $f(x) = 0$ has no multiple root.

2. Solve $x^4 - 6x^3 + 5x^2 + 24x - 36 = 0$ by first constructing a Sturm sequence for the given polynomial.

3. If it is known that the equation $x^4 + 2x^3 + 3x^2 + 2x + 1 = 0$ has no real root, what do you expect for $f_m(x)$? Prove your assertion and solve the given equation.

4. Let $f(x)$ be a polynomial which has no multiple roots.
   (a) For $0 \le k \le m-1$, prove that $f_k(x)$ and $f_{k+1}(x)$ have no common roots.
   (b) For $1 \le k \le m-1$, prove that if $f_k(\alpha) = 0$ for some real number $\alpha$, then $f_{k-1}(\alpha) = -f_{k+1}(\alpha)$.

## 9.2 Sturm's theorem

We have seen in the last section that from a given polynomial $f(x)$ of $\mathbf{R}[x]$ we can derive a Sturm sequence $f_0(x), f_1(x), \ldots, f_m(x)$ by a process of alternate divisions:

$$\begin{aligned}
&f_0(x) = f(x)\\
&f_1(x) = f_1(x)\\
&\cdots\cdots\cdots\\
&f_{k-1}(x) = f_k(x)q_k(x) - f_{k+1}(x)\\
&\cdots\cdots\cdots\\
&f_{m-1}(x) = f_m(x)q_m(x) \ .
\end{aligned}$$

Now every value $c$ of $x$ will give rise to a sequence of values of these functions:

$$f_0(c), f_1(c), \ldots, f_m(c) \ .$$

However for the purpose of isolating the real roots of the equation $f(x) = 0$, our sole interest in this sequence of values lies in their signs and to be more precise, in the number of variations of consecutive signs. Take for instance the sequence of Sturm functions

$$f_0(x) = x^3 + 4x^2 - 8, f_1(x) = 3x^2 + 8x, f_2(x) = \frac{32}{9}x + 8, f_3(x) = \frac{45}{16}$$

of the polynomial $f(x) = x^3 + 4x^2 - 8$ in Example 9.1.1. At $x = 0$ this sequence of Sturm functions yields the following signs,

| | $f_0(x)$ | $f_1(x)$ | $f_2(x)$ | $f_3(x)$ |
|---|---|---|---|---|
| $x = 0$ | $-$ | $0$ | $+$ | $+$ |

Disregarding the zero in the second place, we count 1 variation of signs from $f_0(0) = -$ to $f_2(0) = +$. We shall denote by $V_0$ this number of variations. Here the subscript 0 indicates the value of $x$ at which the counting takes place. Thus $V_0 = 1$. At $x = 1$, we obtain

| | $f_0(x)$ | $f_1(x)$ | $f_2(x)$ | $f_3(x)$ |
|---|---|---|---|---|
| $x = 1$ | $-$ | $+$ | $+$ | $+$ |

yielding also 1 variation: $V_1 = 1$. Similarly at $x = 2$ we have

| | $f_0(x)$ | $f_1(x)$ | $f_2(x)$ | $f_3(x)$ |
|---|---|---|---|---|
| $x = 2$ | $+$ | $+$ | $+$ | $+$ |

yielding no variation: $V_2 = 0$.

We see from the above that there is a difference of 1 between $V_1$ and $V_2$ and there is no difference between $V_0$ and $V_1$. Now according to Sturm's theorem, which we shall state presently, the number $1 = V_1 - V_2$ is precisely the number of roots of $f(x) = 0$ lying between $x = 1$ and $x = 2$, and the number $0 = V_0 - V_1$ is precisely the number of roots $f(x) = 0$ lying between $x = 0$ and $x = 1$. Further tests on the same sequence of Sturm functions will yield the following table:

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| 2 | $+$ | $+$ | $+$ | $+$ | 0 |
| 1 | $-$ | $+$ | $+$ | $+$ | 1 |
| 0 | $-$ | 0 | $+$ | $+$ | 1 |
| $-1$ | $-$ | $-$ | $+$ | $+$ | 1 |
| $-2$ | 0 | $-$ | $+$ | $+$ | |
| $-3$ | $+$ | $+$ | $-$ | $+$ | 2 |
| $-4$ | $-$ | $+$ | $-$ | $+$ | 3 |

The row corresponding to $x = -2$ begins with a 0. This means that $-2$ is a root of the equation $f(x)$. For the purpose of counting this row is to be disregarded; hence a blank at the corresponding place in the last column. The difference $V_{-3} - V_{-1} = 1$ is also accounted for by the presence of the root at $x = -2$. Since $V_{-4} - V_{-3} = 1$, there should be also a real root at the interval between $-4$ and $-3$.

The given equation $x^3 - 4x^2 - 8 = 0$ is a simple cubic equation which can be solved easily. We find $x^3 - 4x^2 - 8 = (x+2)(x^2+2x-4) = (x+2)(x+1+\sqrt{5})(x+1-\sqrt{5})$. It does indeed have a root $-1+\sqrt{5}$ between 1 and 2, a root $-2$ between $-3$ and $-1$, and a root $-1-\sqrt{5}$ between $-4$ and $-3$.

9.2.1 STURM'S THEOREM. *Let $f(x) = 0$ be a polynomial equation and let $f_0(x), f_1(x), \ldots, f_m(x)$ be a sequence of Sturm functions of the polynomial $f(x)$. Then for any two real numbers $a < b$, neither of which is a root*

*of $f(x) = 0$, the number of distinct roots of $f(x) = 0$ lying between $a$ and $b$ equals the difference $V_a - V_b$ between the variations of the signs of the Sturm functions at $x = a$ and $x = b$.*

For the time being, we are far more interested in the applications of Sturm's theorem than in its proof of validity. Since the very lengthy proof is based on several lemmas and a complicated classification of cases, we propose to put it in the appendix at the end of the book, so as not to impede our progress.

The application of Sturm's theorem consists of three separate parts. The first part is the derivation of a series of Sturm functions. The second is to set up a table of signs and to count the number $V_x$ of variations. The third part is to identify the intervals $(a, b)$ at which $V_a$ and $V_b$ have a non-zero difference. Clearly the last two parts, being straight-forward, offer no undue difficulty. On the other hand the alternate divisions of the first part could become very laborious. It is therefore important to pay attention to possible simplifications, so as to diminish labour.

Firstly in order to avoid fractions in the division, we may multiply any one of the functions $f_k(x)$ by a *positive* constant before dividing it by the next $f_{k+1}(x)$. Clearly this will not affect the table of signs.

Secondly, before we use $f_{k+1}(x)$ as the divisor to divide the preceding $f_k(x)$, we may remove from $f_{k+1}(x)$ any factor $g(x)$ which is either a positive constant or a polynomial in $x$ that has positive functional values for all values of $x$. For example, we may replace $f_{k+1}(x) = g(x)h_{k+1}(x)$ by $h_{k+1}(x)$ as the $(k+1)$-th Sturm function if $g(x)$ is of the form, $7$, $x^2 + 2$, $x^2 + x + 1$, $2x^6 - 4x + 5$, $x^4 + x^2 + 3$, etc. We shall see in the appendix that this removal which gives rise to a different series of functions will not affect the final result of the table of signs.

Let us try out with some examples.

9.2.2 EXAMPLE. *For $f(x) = x^5 + 3x^3 + 5x - 10$, we get $f_1(x) = f'(x) = 5x^4 + 9x^2 + 5$ which is always positive. Therefore we may take $f_1(x) = 1$*

*and terminate the process. Moreover, since $f'(x) = 0$ has no real root, we conclude that all real roots of $f(x) = 0$ are simple, if exist. Now we may apply Sturm's theorem to the proposed equation. From the table below we see that the equation $x^5 + 3x^3 + 5x - 10 = 0$ has only one real root which is positive.*

| $x$ | $f_0$ | $f_1$ | $V_x$ |
|---|---|---|---|
| $\infty$ | $+$ | $+$ | 0 |
| 0 | $-$ | $+$ | 1 |
| $-\infty$ | $-$ | $+$ | 1 |

*Further test yields*

| $x$ | $f_0$ | $f_1$ | $V_x$ |
|---|---|---|---|
| 2 | $+$ | $+$ | 0 |
| 1 | $-$ | $+$ | 1 |

*Thus $f(x) = 0$ has one simple root in the interval $(1, 2)$ and four imaginary roots.*

A further test would narrow the interval into $(1, 5/4)$. However this is not necessary because in the next chapter we shall have better ways to approximate the root. The observant reader would have noticed that the equation of the example fits the description of Corollary 8.3.4. Therefore it has at least one positive root, and one of the roots can be easily located in the interval $(0, 2)$ because $f(0) = -$ and $f(2) = +$. But this is all that we can obtain from Bolzano's theorem, because that theorem can give no more information on the existence of further positive roots or negative roots. On the other hand Sturm's theorem gives us the precise and full information that it has no multiple real root and its only real root lies between 1 and 2.

9.2.3 EXAMPLE. *Use Sturm's theorem to analyse the equation*

$$2x^4 - 13x^2 - 10x - 19 = 0 .$$

SOLUTION: Let $f_0(x) = f(x) = 2x^4 - 13x^2 - 10x - 19$. Then dividing $f'(x)$ by 2 we may take $f_1(x) = 4x^3 - 13x - 5$. It follows from

$$2f_0(x) = xf_1(x) - (13x^2 + 15x + 38)$$

that we may take $f_2(x) = 1$, since the negative remainder $g(x) = 13x^2 + 15x + 38$ is a quadratic polynomial with a positive leading coefficient and a negative discriminant. Now $\text{HCF}(f(x), f'(x)) = \text{HCF}(f'(x), g(x))$ and $g(x) = 0$ has no real root; therefore $f(x) = 0$ has no multiple real root. A quick test

| $x$ | $f_0$ | $f_1$ | $f_2$ | $V_x$ |
|---|---|---|---|---|
| $\infty$ | $+$ | $+$ | $+$ | 0 |
| 0 | $-$ | $-$ | $+$ | 1 |
| $-\infty$ | $+$ | $-$ | $+$ | 2 |

shows that the equation has one positive and one negative root, both being simple. A better separation is achieved by the following table:

| $x$ | $f_0$ | $f_1$ | $f_2$ | $V_x$ |
|---|---|---|---|---|
| 4 | $+$ | $+$ | $+$ | 0 |
| 3 | $-$ | $+$ | $+$ | 1 |
| $-2$ | $-$ | $-$ | $+$ | 1 |
| $-3$ | $+$ | $-$ | $+$ | 2 |

Thus the two real roots of the equation lie in the interval $(3, 4)$ and $(-3, -2)$.

9.2.4 EXAMPLE. *Analyse the equation*

$$x^3 + 11x^2 - 102x + 181 = 0 \ .$$

SOLUTION: Let $f_0(x) = f(x) = x^3 + 11x^2 - 102x + 181$. Then $f_1(x) = f'(x) = 3x^2 + 22x - 102$. Division of $9f_0(x)$ by $f_1(x)$ gives

$$9f_0(x) = (3x + 11)f_1(x) - (854x - 2751) \ .$$

Put $f_2(x) = (854x - 2751)/854$. By the remainder theorem, $f_1(x) = q_2(x)f_2(x) + f_1(\frac{2751}{854})$; we may put $f_3(x) = -f_1(\frac{2751}{854}) = +$. Thus all roots of $f(x) = 0$ are simple. The table

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | $+$ | $+$ | $+$ | $+$ | 0 |
| 4 | $+$ | $+$ | $+$ | $+$ | 0 |
| 3 | $+$ | $-$ | $-$ | $+$ | 2 |
| $-17$ | $+$ | $+$ | $-$ | $+$ | 2 |
| $-18$ | $-$ | $+$ | $-$ | $+$ | 3 |
| $-\infty$ | $-$ | $+$ | $-$ | $+$ | 3 |

shows that there are two simple roots in $(3,4)$ and one simple root in $(-18,-17)$.

9.2.5 EXAMPLE. *Analyse the equation*

$$x^3 - 2x - 5 = 0 .$$

SOLUTION: Let $f_0(x) = f(x) = x^3 - 2x - 5$. Then $f_1(x) = f'(x) = 3x^2 - 2$, and $3f_0(x) = xf_1(x) - (4x + 15)$. Therefore we can put $f_2(x) = x + \frac{15}{4}$. Since $f_1(-\frac{15}{4}) > 0$ we obtain $f_3(x) = -$. The table

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | + | + | + | − | 1 |
| 0 | − | − | + | − | 2 |
| $-\infty$ | − | + | − | − | 2 |

shows that the equation has one real root which is positive. Choosing 2 and 3 in the inteval $(0,\infty)$ we get

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| 3 | + | + | + | − | 1 |
| 2 | − | + | + | − | 2 |

Therefore the equation has one simple real root between 2 and 3 and two distinct imaginary roots.

The observant reader must have noticed that in all the previous four examples the last functions of the Sturm sequences are either non-zero constants or polynomials which have no real roots. Therefore the equations in questions have no multiple real roots. The same method can be applied to equations with multiple real roots without modification. In fact in all cases, *the difference $V_a - V_b$ is the number of real roots between $a$ and $b$, each multiple root counted only once.* Thus $V_a - V_b$ is the number of *distinct* real roots between $a$ and $b$.

9.2.6 EXAMPLE. *Analyse the equation*

$$x^4 - 5x^3 + 9x^2 - 7x + 2 = 0 .$$

SOLUTION: For $f(x) = x^4 - 5x^3 + 9x^2 - 7x + 2$, we find that

$$\begin{aligned} 4^3 f(x) &= (16x - 20) f'(x) - 12(x^2 - 2x + 1) \\ f'(x) &= (4x - 7)(x^2 - 2x + 1) \,. \end{aligned}$$

Therefore

$$\begin{aligned} f_0(x) &= x^4 - 5x^3 + 9x^2 - 7x + 2 \\ f_1(x) &= 4x^3 - 15x^2 + 18x - 7 \\ f_2(x) &= x^2 - 2x + 1 \end{aligned}$$

form a Sturm sequence. A preliminary test shows that

| $x$ | $f_0$ | $f_1$ | $f_2$ | $V_x$ |
|---|---|---|---|---|
| $\infty$ | + | + | + | 0 |
| 0 | + | − | + | 2 |
| $-\infty$ | + | − | + | 2 |

According to Sturm's theorem the equation $f(x) = 0$ should have 2 distinct positive roots. Indeed we have $f_2(x) = \text{HCF}(f(x), f'(x)) = x^2 - 2x + 1 = (x-1)^2$ and $f(x) = (x-1)^3(x-2)$. Therefore the equation has a triple root at $x = 1$ and a simple root at $x = 2$. We observe that since the table gives the number of distinct roots, there is no uncertainty as to whether some root has been missed out.

9.2.7 EXAMPLE. *Separate the roots of the equation*

$$x^4 - 3x^3 + x^2 + 4 = 0 \,.$$

SOLUTION: For $f(x) = x^4 - 3x^3 + x^2 + 4 = 0$ we find that

$$\begin{aligned} 4^3 f(x) &= (16x - 12) f'(x) - 4(19x^2 - 6x - 64) \\ 19^2 f'(x) &= (76x - 147)(19x^2 - 6x - 64) - 4704(-x + 2) \\ 19x^2 - 6x - 64 &= (-19x - 32)(-x + 2) \end{aligned}$$

yielding the following sequence of Sturm functions:

$$f_0(x) = x^4 - 3x^3 + x^2 + 4$$

$$f_1(x) = 4x^3 - 9x^2 + 2x$$
$$f_2(x) = 19x^2 - 6x - 64$$
$$f_3(c) = -x + 2 \ .$$

Thus HCF$(f(x), f'(x)) = -x + 2$. Hence 2 is a double root of the equation. The degree of $f(x)$ being 4, we still have to see if there are other real roots. From the table

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | $+$ | $+$ | $+$ | $-$ | 1 |
| $-\infty$ | $+$ | $-$ | $+$ | $+$ | 2 |

we see that $f(x) = 0$ has only one real root which is the double root 2, already found. Therefore the remaining roots of the equation are imaginary.

We take note that the graph of $f(x) = x^4 - 3x^3 + x^2 + 4 = 0$ lies entirely in the upper half-plane. Therefore it would be very difficult for us to find out the distribution of its real roots by Bolzano's method alone.

To conclude this section we consider again the general cubic equation with vanishing quadratic term in the light of Sturm's theorem.

9.2.8 EXAMPLE. *Given a cubic equation of the form*

$$x^3 + px + q = 0 \ .$$

*Find the condition on the coefficients $p$ and $q$ so that the roots should be all real and distinct.*

SOLUTION: We put $f_1(x) = x^2 + \frac{1}{3}p$. Then

$$x^3 + px + q = x(x^2 + \frac{1}{3}p) + (\frac{2}{3}px + q)$$
$$\frac{4}{9}p^2(x^2 + \frac{1}{3}p) = (-\frac{2}{3}px + q)(-\frac{2}{3}px - q) + (\frac{4}{27}p^3 + q^2) \ .$$

Therefore

$$\begin{aligned} f_0(x) &= x^3 + px + q \\ f_1(x) &= 3x^2 + p \\ f_2(x) &= -(2px + 3q) \\ f_3(x) &= -(4p^3 + 27q^2) \end{aligned}$$

form a Sturm sequence for $f(x) = x^3 + px + q$. Clearly the necessary and sufficient condition for $f(x) = 0$ to have three distinct roots is that $f_3(x)$ is a non-zero constant, i.e. $4p^3 + 27q^2 \neq 0$. Consequently we shall have to investigate the following two cases.

Case (a) $4p^3 + 27q^2 > 0$. If $p > 0$ then Sturm's method would give

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | + | + | − | − | 1 |
| $-\infty$ | − | + | + | − | 2 |

If $p < 0$ then Sturm's method would give

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | + | + | + | − | 1 |
| $-\infty$ | − | + | − | − | 2 |

Therefore in either case we would have only one real root instead of three.

Case (b) $4p^3 + 27q^2 < 0$. Then $p < 0$ and

| $x$ | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $V_x$ |
|---|---|---|---|---|---|
| $\infty$ | + | + | + | + | 0 |
| $-\infty$ | − | + | − | + | 3 |

Therefore $f(x) = 0$ has three distinct real roots.

Thus a necessary and sufficient condition for $x^3 + 3p + q = 0$ to have three distinct real roots is that $4p^3 + 27q^2 < 0$.

Since the imaginary roots of an equation with real coefficients always appear in conjugate pairs, the equation $x^3 + px + q = 0$ cannot have multiple imaginary root. Hence case (a) in the proof above yields that the equation has a single real root and a pair of conjugate imaginary roots if and only if $4p^3 + 27q^2 > 0$. The remaining case

where $4p^3 + 27q^2 = 0$, say case (c), would therefore hold if and only if the equation has three real roots, at least two being equal. Therefore the cases (a), (b) and (c) are exactly the three cases of Theorem 4.3.8 where the discriminant $\Delta = -4p^3 - 27q^2$ of the equation is negative, positive or zero respectively.

## EXERCISE 9B

In Questions 1 to 6, analyse the equation using Sturm's theorem.

1. $x^3 + x^2 - 2x - 2 = 0$.
2. $12x^3 - 32x^2 + 25x - 6 = 0$.
3. $x^4 + 4x^3 + 6x^2 + 20x + 5 = 0$.
4. $x^4 - 2x^3 + x^2 - 4x - 2 = 0$.
5. $9x^4 - 12x^3 + 13x^2 - 12x + 4 = 0$.
6. $x^4 - 8x^3 + 19x^2 - 12x - 4 = 0$.
7. Use Sturm's theorem to prove that the roots of a real quadratic polynomial $ax^2 + bx + c$ are real and distinct if and only if $b^2 - 4ac > 0$.

## 9.3 Fourier's theorem

Though Sturm's method allows us to separate all the real roots from each other it has a disadvantage in that the Sturm functions had to be founded by a series of laborious division algorithms. Indeed for an equation of degree higher than 5, the coefficients of $f_2(x), f_3(x), \dots$ may be quite unmanageable. Before the discovery of Sturm, the French mathematician Joseph Fourier (1768–1830) had found a method of separation using only the derivatives.

Let $f(x)$ be a polynomial of degree $n$ and let

$$f'(x), f''(x), \dots, f^{(k)}(x), \dots, f^{(n)}(x)$$

be its successive derivatives. For any real number $a$ which is not a root of the equation $f(x) = 0$, we denote by $W_a$ the number of variations of signs of the values

$$f(a), f'(a), \ldots, f^{(n-1)}(a), f^{(n)}(a)$$

after the vanishing terms are deleted.

9.3.1 FOURIER'S THEOREM. *Let $a < b$ be two real numbers, neither being a root of $f(x) = 0$. Then $W_a - W_b$ is either the number of roots of $f(x) = 0$ in the interval $(a, b)$ or exceeds the number of those roots by an even integer. A root of multiplicity $m$ is counted as $m$ roots here.*

The theorem is also named after the French physician F.D. Budan, a contemporary of Fourier though he did not actually prove it. We shall sketch a proof of the theorem in the appendix. Meanwhile we take note that, except in the case where $W_a - W_b$ is 1 or 0, there is no way to tell which of the quantities $W_a - W_b$, $W_a - W_b - 2$, $W_a - W_b - 4, \ldots$ is the correct number of roots of $f(x) = 0$ in $(a, b)$. In particular if $W_a - W_b$ is even and non-zero, even the existence of roots in $(a, b)$ is uncertain. On the other hand, if $W_a - W_b$ is odd, then we know that at least one root lies between $a$ and $b$. At any rate, the quantity $W_a - W_b$ gives the maximum number of roots in the interval.

9.3.2 EXAMPLE. *Apply Fourier's theorem to the equation*

$$x^3 - 7x - 7 = 0 .$$

SOLUTION: We have $f(x) = x^3 - 7x - 7$, $f'(x) = 3x^2 - 7$, $f''(x) = 6x$, $f'''(x) = 6$.

Since the leading coefficient and the constant term have opposite signs, we can conclude by Corollary 8.3.4 that $f(x) = 0$ has at least one positive root. We now apply the test to see if there is just one positive root.

| $x$ | $f$ | $f'$ | $f''$ | $f'''$ | $W$ |
|---|---|---|---|---|---|
| $\infty$ | + | + | + | + | 0 |
| 4 | + | + | + | + | 0 |
| 3 | − | + | + | + | 1 |
| 0 | − | − | 0 | + | 1 |

The test confirms that there is exactly one positive root and this root lies

between 3 and 4. For the negative values of $x$ no information can be obtained from Corollary 8.3.3, but we find

| $x$ | $f$ | $f'$ | $f''$ | $f'''$ | $W$ |
|---|---|---|---|---|---|
| $0$ | $-$ | $-$ | $0$ | $+$ | $1$ |
| $-1$ | $-$ | $-$ | $-$ | $+$ | $1$ |
| $-2$ | $-$ | $+$ | $-$ | $+$ | $3$ |
| $-\infty$ | $-$ | $+$ | $-$ | $+$ | $3$ |

Thus by Fourier's theorem 9.3.1 there are two roots or there is no root between $-2$ and $-1$. The former case will occur if a sequence of signs $+\ +\ -\ +$ or $+\ -\ -\ +$ is obtained for some value between $-2$ and $-1$ since $f''$ and $f'''$ will remain $-$ and $+$ respectively for any such value. Now $f'(x) = 3x^2 - 7$ has a root at $-\sqrt{7/3} = -1.52\ldots$ . For $x = -1.6$ we do get

| $x$ | $f$ | $f'$ | $f''$ | $f'''$ | $V$ |
|---|---|---|---|---|---|
| $-1.6$ | $+$ | $+$ | $-$ | $+$ | $2$ |

Therefore we have one more root between $-2$ and $-1.6$ and another root between $-1.6$ and $-1$.

## Exercise 9C

In Questions 1 to 6, analyse the equations by Fourier's theorem.

1. $x^3 + x^2 - 2x - 2 = 0$.
2. $12x^3 - 32x^2 + 25x - 6 = 0$.
3. $x^4 + 4x^3 + 6x^2 + 20x + 5 = 0$.
4. $x^4 - 2x^3 + x^2 - 4x - 2 = 0$.
5. $9x^4 - 12x^3 + 13x^2 - 12x + 4 = 0$.
6. $x^4 - 8x^3 + 19x^2 - 12x - 4 = 0$.

### 9.4 Descartes' rule of signs

Long before the discoveries by Fourier and Sturm in the nineteenth century, the French philosopher and mathematician René

Descartes (1596–1650), who is usually credited for the discovery of analytic geometry, gave a rule of signs for determining the positive and negative roots of an equation. This is contained in his celebrated book *La géométrie* published in 1637. He described the rule in only one single instance using a quartic equation. The rule was extended by Isaac Newton (1642–1727) and later proved by Jean Paul de Gua (1713–1785). Finally it is given in the following form and proved by Carl Friedrich Gauss (1777–1855).

9.4.1 DESCARTES' RULE OF SIGNS. *The number of positive roots of an equation $f(x) = 0$ either equals the number $V$ of variations of the signs in the series $a_n, a_{n-1}, \ldots, a_1, a_0$ of the coefficients of $f(x)$ or less than $V$ by an even integer. A root of multiplicity $m$ is here counted as $m$ roots.*

PROOF: Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $V$ be the number of variations of signs of the series $a_n, a_{n-1}, \ldots, a_1, a_0$. Consider first the case where $a_0 \neq 0$. Then for a sufficiently large positive value $b$, all the functional values $f(b), f'(b), \ldots, f^{(n-1)}(b), f^{(n)}(b)$ have the same sign as the leading coefficient $a_n$. Therefore $W_b = W_\infty = 0$. On the other hand, the signs of the series $f(0), f'(0), \ldots, f^{(n-1)}(0), f^{(n)}(0)$ are the same as those of $a_0, a_1, \ldots, a_{n-1}, a_n$. Therefore $W_0 = V$ and $W_0 - W_\infty = V$. Hence Descartes' rule follows. If $a_0 = 0$, then we can write $f(x) = x^m g(x)$ where $g(x)$ has a non-vanishing constant term. Now $g(x) = 0$ and $f(x) = 0$ have the same number of positive roots. Also the number of variations of signs of the coefficients of $g(x)$ is the same as that of $f(x)$. Therefore the rule also holds for the case where $a_0 = 0$.

9.4.2 COROLLARY. *The number of negative roots of an equation $f(x) = 0$ is either the number of variations of signs of the coefficients of $f(-x)$ or is less than that number by an even integer.*

9.4.3 EXAMPLE. *The equation $x^3 - 3x + 2 = 0$ has one negative root and two equal positive roots.*

SOLUTION: Let $f(x) = x^3 - 3x + 2$. Then $f(-x) = -x^3 + 3x + 2$ and the signs of its coefficients are $-\ +\ +$. Therefore $V = 1$ and $f(x) = 0$ has

one negative root. The signs of the coefficients of $f(x)$ are $+\ -\ +$. Since $V = 2$ we do not have conclusive information from Descartes' rule. However $f(1) = f'(1) = 0$. Therefore 1 is a positive double root of $f(x) = 0$.

9.4.4 EXAMPLE. *The equation $x^3 + a^2x + b^2 = 0$ has two imaginary roots if $b \neq 0$.*

SOLUTION: By Descartes' rule the cubic equation has one negative root and no positive root. Therefore the two remaining must be imaginary and distinct. Alternatively we may examine the discriminant

$$\Delta = -4p^3 - 27q^2$$

of the cubic equation

$$y^3 + py + q = 0$$

as given in Definition 4.3.7. For the present equation, we get $\Delta < 0$. Therefore the equation has only one real root and two imaginary roots.

## EXERCISE 9D

1. Determine the possible number of positive and negative roots of the following equations using Descartes' rule of signs.
   (a) $x^3 + x^2 - 2x - 2 = 0$.
   (b) $x^4 + 4x^3 + 6x^2 + 20x + 5 = 0$.
   (c) $12x^3 - 32x^2 + 25x - 6 = 0$.
   (d) $x^4 - 2x^3 + x^2 - 4x - 2 = 0$.
   (e) $9x^4 - 12x^3 + 13x^2 - 12x + 4 = 0$.
   (f) $x^4 - 8x^3 + 19x^2 - 12x - 4 = 0$.
2. Given

$$x^7 - 4x^6 + 5x^5 - 4x + 3 = 0 \qquad (*)$$

and

$$x^7 + 4x^6 - 5x^5 - 4x + 3 = 0 \qquad (**)\,.$$

   (a) Prove that $(*)$ and $(**)$ cannot have more than four and two positive roots respectively.

(b) Show that (*) has at least two complex roots and that (**) has at least four complex roots.

3. Prove Corollary 9.4.2.

4. Given two real numbers $p$, $q$, show that the equation

$$x^3 + p^2x + q = 0$$

must have complex roots for all values of $p \neq 0$ and $q$.

5. Prove that the real polynomial $x^4 + a^2x^2 + b^2x - c^2 = 0$, $c \neq 0$, has exactly 2 real roots.

6. Show that, for any natural number $n$,

$$x^{2n} + x^{2n-2} + x^{2n-4} + \cdots + x^2 + 1 = 0$$

has no real roots.

7. Let $f(x) = x^5 + 2x^3 - x^2 + x - 1$.

   (a) Find a monic real polynomial $g(x)$ whose roots are the squares of the roots of $f(x)$.

   [Hint: Consider $f(x) \cdot f(-x)$.]

   (b) Use Descartes' rule of signs to study the non-negative real roots of $g(x)$. Hence conclude that $f(x)$ has four complex roots.

8. Let $f(x)$ be a polynomial of degree $n$ with $n$ real roots. If $h$ is the number of change sign of the coefficients of $f(x)$ and $n - h$ is the number of change sign of the coefficients of $f(-x)$, show that $f(x) = 0$ has $h$ positive roots and $n - h$ negative roots.

9. Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a non-zero polynomial with $n > 1$.

   (a) (i) Find $g(x)$ if $g(x^2) = f(x)f(-x)$.

   (ii) Consider the coefficient of $x$ in $g(x)$, show that if $a_1^2 < 2a_2a_0$, then $f(x)$ has complex roots.

   (b) For $1 \leq k \leq n - 1$, find $f^{(n-k-1)}(x)$ and hence show that if $(a_{n-k})^2 < a_{n-k+1} \cdot a_{n-k-1}$, then $f(x)$ has complex roots.

CHAPTER TEN

# APPROXIMATION TO REAL ROOTS

We recall that given an equation of degree less than five, the exact values of its roots can be written as expressions that involve only rational operations and root extractions on the coefficients. It is also known that such expressions of roots are not generally available for an equation of higher degree. Therefore, for such equations, we shall have to use numerical methods that would only give approximate decimal values to the real roots. An approximate value is always inferior to the exact value, but for many practical purposes, we only need good approximations. In this chapter we shall learn two iterative processes that can furnish approximations to roots to any desired degree of accuracy.

## 10.1 Newton-Raphson method

In 1669 Newton published a treatise entitled *De Analysi per Aequationes Numero Terminorum Infinitas* in which he explained a method of approximating roots of numerical equations by working out one example, namely the cubic equation $x^3 - 2x - 5 = 0$.

10.1.1 NEWTON'S EXAMPLE. *The cubic equation*

$$x^3 - 2x - 5 = 0$$

*has a root between* 2 *and* 3. *Find approximate values to this root.*

SOLUTION: The discriminant of the equation $x^3 - 2x - 5 = 0$ is negative; hence by 4.3.8 it has only one real root $r$. With $f(2) = -1$ and $f(3) = 16$, we locate $r$ between 2 and 3. Newton's method consists of a series of successive approximations. To begin we take $a_1 = 2$ as the first rough approximation

to $r$, and proceed to find the next approximation $a_2 = a_1 + h_1$ with a small correcting term $h_1$. Thus a decimal value of $h_1$ is to be found so that

$$\begin{aligned} f(a_2) &= f(2+h_1) \\ &= -5 - 2(2+h_1) + (2+h_1)^3 \\ &= (-1 + 10h_1) + (6{h_1}^2 + {h_1}^3) \end{aligned}$$

would be close to 0. Obviously it would simplify the matter if we take $h_1 = -1/10 = 0.1$. Then the value in the first bracket would be zero and the value in the second bracket which then equals $f(a_2) = f(2.1)$ would be small because it only contains the quadratic and the cubic terms of the decimal 0.1. Indeed with $f(2.1) = 0.061$ we could accept $a_2 = a_1 + h_1 = 2.1$ as a better approximation to $r$ than $a_1 = 2$ with $f(2) = -1$.

Should a closer approximation than 2.1 be needed, we proceed to find $a_3 = a_1 + h_1 + h_2$ with a still smaller correcting term $h_2$. To find $h_2$, we try to make

$$\begin{aligned} f(a_3) &= f(2 + (0.1 + h_2)) \\ &= -1 + 10(0.1+h_2) + 6(0.1+h_2)^2 + (0.1+h_2)^3 \\ &= (0.061 + 11.23h_2) + (6.3{h_2}^2 + {h_2}^3) \end{aligned}$$

still closer to 0 than 0.061. Similarly we choose $h_2 = -0.061/11.23 \approx -0.0054$ to make the first bracket vanish and consequently diminish $f(2.1)$ to

$$f(2.0946) = f(2.1 - 0.0054) \approx 0.0005415\ .$$

With this improvement, we could accept $a_3 = 2.0946$ as the third approximation to $r$.

Thus we have obtained three approximate values to the real root $r$ of $x^3 - 2x - 5 = 0$ in

$$a_1 = 2, \quad a_2 = 2.10, \quad a_3 = 2.0946$$

with

$$f(a_1) = -1, \quad f(a_2) = 0.061, \quad f(a_3) \approx 0.0005415\ .$$

Further approximations $a_4, a_5, \ldots$ may be found similarly.

We may describe Newton's method in general terms as follows. Given is an equation $f(x) = 0$ (e.g., $f(x) = x^3 - 2x^2 - 5 = 0$) together with an approximation $a$ (e.g., $a = 2$) to one of its root $r$. Then we consider a new equation

$$g_0(h) = f(a+h) = 0$$

in the unknown $h$ which has a root at $r - a$. Neglecting all terms in $h^2, h^3$, etc., we find an approximation $h_1$ (e.g., $h_1 = 0.1$) to the root $r - a$. Then $a + h_1$ (e.g., 2.1) will be a new approximation to the root $r$ of $f(x) = 0$. If a better one is desired, the next step takes us to yet another new equation

$$g_1(h') = g_0(h_1 + h') = 0$$

in the unknown $h'$. This equation has a root in $r - a - h_1$. Again an approximation $h_2$ (e.g., $h_2 = -0.0054$) to this root is found after neglecting the higher terms. Thus an improved approximation $a + h_1 + h_2$ (e.g., 2.0946) is obtained. To find the next approximation we consider the equation

$$g_2(h'') = g_1(h_2 + h'') = 0$$

in the unknown $h''$ which has a root $r - a - h_1 - h_2$. Similarly an approximation $h_3$ (e.g., $-0.00004852$) to this root of $g_2(h'') = 0$ is found giving rise to the next approximation $a + h_1 + h_2 + h_3$ (e.g., 2.09455148) to the root $r$ of $f(x) = 0$. The process is to be terminated as soon as the desired accuracy of the approximation is attained.

In 1690, Joseph Raphson (1648–1715) proposed a method very similar to Newton's . The only difference is this: Newton derives each correcting term $h_1, h_2, \ldots$ from a new equation $g_0(x) = 0$, $g_1(x) = 0, \ldots$, while Raphson finds it each time by substitution in the original equation $f(x) = 0$. This method is now called the *Newton-Raphson method*. We shall now describe this method in general terms.

Let $f(x) = 0$ be an equation of degree $n$, and $a$ be any one approximation to a true root $r$ of $f(x) = 0$. Then we proceed to find a correcting term $h$ to $a$ so that $a + h$ is closer to $r$ than $a$ is. After a substitution of $a + h$ for $x$ in $f(x) = 0$, Taylor's expansion gives

$$f(a+h) = \{f(a) + f'(a)h\} + \{f''(a)\frac{h^2}{2} + \cdots + f^{(n)}(a)\frac{h^n}{n!}\} = 0 \ .$$

As $a$ is close to $r$, we may take for $h$ a value which is small enough for us to neglect the value of the second bracket in the above expansion. Thus we obtain from $f(a) + f'(a)h \approx 0$ a correcting term

$$h = -\frac{f(a)}{f'(a)} .$$

For the next step we regard $a' = a+h$ as a new approximation to $r$ and proceed to find a correcting term $h'$ to $a'$. Using the same argument as in the previous step, we substitute $a'+h'$ for $x$ in $f(x) = 0$ to obtain

$$f(a' + h') = \{f(a') + f'(a')h'\} + \text{higher terms in } h' .$$

As the correcting term to $a'$ we use

$$h' = -\frac{f(a')}{f'(a')}$$

furnishing us with a new approximation $a'' = a' + h'$ to the root $r$ of $f(x) = 0$. A repetition of the process provides the next approximation to $r$, and so on.

10.1.2 EXAMPLE. *Find, correct to four decimal places the root between 1 and 2 of the equation*

$$x^3 + 4x^2 - 7 = 0 .$$

SOLUTION: For $f(x) = x^3 + 4x^2 - 7$ we have $f'(x) = 3x^2 + 8x$. Denoting by $h_i$ the correcting term to the approximate value $a_i$, we get

$$a_1 = 1; \quad h_1 = -\frac{f(a_1)}{f'(a_1)} = -\frac{f(1)}{f'(1)} = \frac{2}{11} = 0.1\cdots$$

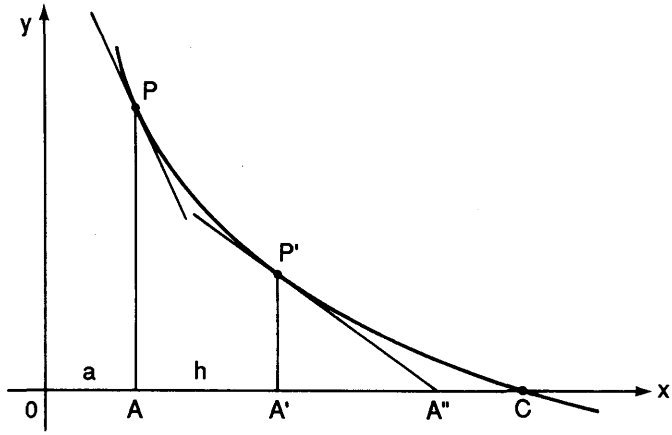
$$a_2 = 1.1; \quad h_2 = -\frac{f(a_2)}{f'(a_2)} = -\frac{f(1.1)}{f'(1.1)} = \frac{0.829}{12.43} = 0.06\cdots$$

$$a_3 = 1.16; \quad h_3 = -\frac{f(a_3)}{f'(a_3)} = -\frac{f(1.16)}{f'(1.16)} = \frac{0.056704}{13.3168} = 0.004\cdots$$

$$a_4 = 1.164; \quad h_4 = -\frac{f(a_4)}{f'(a_4)} = -\frac{f(1.164)}{f'(1.164)} = \frac{0.00317056}{13.376688} = 0.0002\cdots$$

We find $a_5 = a_4 + h_3 = 1.1642$, a root of $f(x)$ correct to four decimal places.

Geometrically the Newton-Raphson method can be explained by the graph:




The curve represents the graph of the polynomial $y = f(x)$ which crosses the $x$-axis at the point $C = (r, 0)$, $r$ being a true root of the equation $f(x) = 0$ to be approximated. $A = (a, 0)$ is the first approximation to $C = (r, 0)$ and $PA'$ is the tangent to the curve at the point $P$. Then $f'(a) = \tan \angle CA'P = -f(a)/h$. Therefore $A' = (a + h, 0)$ is the second approximation to $C = (r, 0)$.

From the graph we also see that the method would not be efficient if the point $P = (a, f(a))$ is in close proximity to a bend point or an inflexion point of the curve. On the other hand, we see from its derivation that the method is applicable to any function for which Taylor's formula holds. Since such functions include most functions studied in the calculus, the method has a very wide scope of application. However for polynomials, the Qin-Horner method which we shall study in the next section proves to be more superior.

## EXERCISE 10A

1. Using Newton-Raphson method, find, correct to three decimal places, the real root of each of the following equations in the specified intervals.
   (a) $x^3 + x^2 - 2x - 2 = 0$ in (1, 2).
   (b) $x^4 - 2x^3 + x^2 - 4x - 2 = 0$ in $(-1, 0)$.
   (c) $x^4 - 8x^3 + 19x^2 - 12x - 4 = 0$ in (4, 5).

2. Find, correct to three decimal places, all the real roots of $7x^3 - 6x^2 + 7x - 6 = 0$.

   [Hint: You should use some results learnt before.]

3. Find, correct to two decimal places, all the positive real roots of $x^4 - 5x^2 + 8x - 8 = 0$.

4. To approximate $\sqrt[p]{k}$, where $k > 0$ and $p$ is a positive integer, by the Newton-Raphson method, we solve $x^p - k = 0$. Show that the correcting term for $a'$ is

$$h = \frac{(a')^p - k}{p(a')^{p-1}} .$$

   Hence find the values of (a) $\sqrt[3]{2}$, and (b) $\sqrt[6]{5}$ corrected to four decimal places.

## 10.2 Qin-Horner method

In this section we shall learn another method of approximation which is commonly named after William George Horner (1786–1837) who published the method in 1819. A similar method was employed by Paolo Ruffini (1765–1822) in 1804. But unknown to them, this same method was used by Jia Xian 賈憲 in the 11-th century to extract roots. Jia's method was then improved and extended by Qin Jiu Shao 秦九韶 (1202–1261) to solve polynomial equations. Qin's method is explained in great detail in his book *Shu Shu Jiu Zhang* 數書九章 published in 1247.

Let us use the cubic equation

$$f(x) = x^3 - 2x - 5 = 0$$

of Newton's example to illustrate its working and point out the essential differences between this method and the method of the last section. By Newton's method the polynomial $f(x)$ is 'transformed' into the polynomial

$$g(h) = f(2 + h) = (2 + h)^3 - 2(2 + h) - 5$$

where 2 is used as an estimated value of the root of $f(x) = 0$ with $h$ as a correcting term. The coefficients of $g(h)$ are calculated by

expanding the binomials on the right-hand side. Thus

$$g(h) = h^3 + 6h^2 + 10h - 1$$

and an approximate value to $h$ is found by solving the equation $g(h) = 0$.

The first difference between the two methods lies in the way in which the coefficients of $g(h)$ are calculated. We shall see that instead of expanding binomials, we can obtain the coefficients of $g(h)$ by a series of synthetic substitutions using the coefficients of $f(x)$ alone. Now it follows from $g(h) = f(2+h)$ that

$$g(x-2) = f(2+(x-2)) = f(x) \ .$$

Therefore the polynomial identity $f(x) = g(x-2)$ can be written as

$$x^3 - 2x - 5 = (x-2)^3 + 6(x-2)^2 + 10(x-2) - 1 \ .$$

The coefficients of $g(h)$ which we seek are the coefficients of $g(x-2)$ on the right-hand side of the identity. But they are precisely the constant terms of the following polynomials:

$$\begin{aligned} g(x-2) = g_0(x-2) &= (x-2)^3 + 6(x-2)^2 + 10(x-2) - 1 \\ g_1(x-2) &= (x-2)^2 + 6(x-2) + 10 \\ g_2(x-2) &= (x-2) + 6 \\ g_3(x-2) &= 1 \ . \end{aligned}$$

Here each $g_{i+1}(x-2)$ is just the quotient of the division of the previous $g_i(x-2)$ by $(x-2)$. Moreover the coefficients of $g(x-2)$ and hence of $g(h)$ are the remainders of these successive divisions.

Carrying out the corresponding successive divisions by $(x-2)$ on the polynomial $f(x)$, we would get

$$\begin{aligned} f(x) = f_0(x) &= (x-2)f_1(x) + f_0(2) = (x-2)(x^2+2x+2) - 1 \\ f_1(x) &= (x-2)f_2(x) + f_1(2) = (x-2)(x+4) + 10 \\ f_2(x) &= (x-2)f_3(x) + f_2(2) = (x-2) + 6 \\ f_3(x) &= (x-2)f_4(x) + f_3(2) = 1 \ . \end{aligned}$$

Because $f_0(x) = g_0(x-2)$ the successive Euclidean algorithm on both sides must yield identitical quotients and remainders:

$$\begin{aligned} f_0(x) &= g_0(x-2) \\ f_1(x) &= g_1(x-2) \text{and } f_0(2) = -1 \\ f_2(x) &= g_2(x-2) \text{and } f_1(2) = 10 \\ f_3(x) &= g_3(x-2) \text{and } f_2(2) = 6 \\ 0 = f_4(x) &= g_4(x-2) \text{and } f_3(2) = 1 \ . \end{aligned}$$

Recall that the coefficients of the quotient and the remainder of a division by $(x-2)$ can be obtained by a synthetic substitution. Therefore we can use the coefficients of $f(x)$ to rewrite the successive divisions into the following schemes:

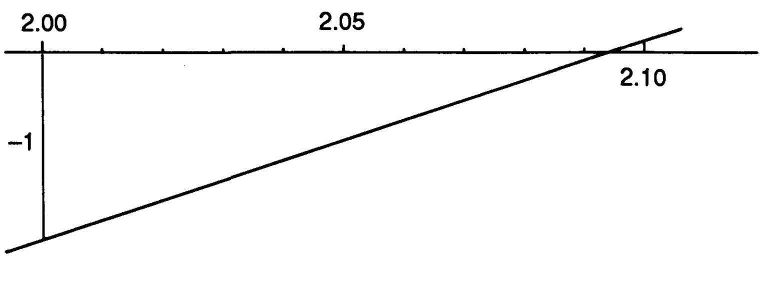
$$\begin{array}{l|cccc l}
\underline{2} & 1 & 0 & -2 & -5 & f_0(x) = x^3 - 2x - 5 \\
 & & 2 & 4 & 4 & \\ \cline{2-5}
 & 1 & 2 & 2 & \underline{-1} & f_1(x) = x^2 + 2x + 2, f_0(2) = -1 \\
\underline{2} & 1 & 2 & 2 & & \\
 & & 2 & 8 & & \\ \cline{2-4}
 & 1 & 4 & \underline{10} & & f_2(x) = x + 4, f_1(2) = 10 \\
\underline{2} & 1 & 4 & & & \\
 & & 2 & & & \\ \cline{2-3}
 & 1 & \underline{6} & & & f_3(x) = 1, f_2(x) = 6 \\
\underline{2} & 1 & & & & \\
 & 0 & & & & \\ \cline{2-2}
 & \underline{1} & & & & f_4(x) = 0, f_3(2) = 1
\end{array}$$

Therefore instead of expanding the binomials in $(2+h)^3 - 2(2+h) - 5$ as suggested by Newton, we can also obtain the coefficients of $g(h)$ by successive synthetic substitutions of the value $x = 2$ in $f(x)$ and the subsequent quotients. In fact, we may amalgamate the synthetic substitutions into one scheme as follows:

$$
\begin{array}{r|rrrr}
2 & 1 & 0 & -2 & -5 \\
  &   & 2 & 4 & 4 \\
\hline
  & 1 & 2 & 2 & -1 \\
  &   & 2 & 8 & \\
\hline
  & 1 & 4 & 10 & \\
  &   & 2 & & \\
\hline
  & 1 & 6 & &
\end{array}
$$

with the coefficients of $g(h) = h^3 + 6h^2 + 10h - 1$ appearing at the lower ends of the columns.

The second difference between the two methods lies in the determination of the value $h$. Newton, after neglecting the quadratic and cubic terms of $g(h) = 0$, used the value $-(\frac{-1}{10}) = 0.1$ for the correcting term $h$. Now the value 0.1 exceeds the true root of $g(h) = 0$ and hence $2 + 0.1 = 2.1$ exceeds the root $r$ of $f(x) = 0$ leading to a negative correcting term at the next step.




Qin and Horner would have used 0.09 for the correcting term $h$, by taking a decimal with a single significant figure less than 0.1. This is to ensure that $2 + 0.09$ lies to the left of the root of $f(x) = 0$.

There is a third difference between the two methods. To obtain the next correcting term $k$ to 2.1, Raphson worked with the polynomial $f(2.1 + k)$ while Qin and Horner, agreeing with Newton, would work with the polynomial $g(0.09 + k)$ and apply the same process as before. Schematically the second step of the method is as follows:

| | | | | |
|---|---|---|---|---|
| 0.09 | 1 | 6.00 | 10.0000 | −1.000000 |
| | | 0.09 | 0.5481 | 0.949329 |
| | 1 | 6.09 | 10.5481 | −0.050671 |
| | | 0.09 | 0.5562 | |
| | 1 | 6.18 | 11.1043 | |
| | | 0.09 | | |
| | 1 | 6.27 | | |

This yields a negative remainder −0.050671 as desired. Moreover we get

$$\begin{aligned} p(k) = g(0.09+k) &= (0.09+k)^3 + 6(0.09+k)^2 + 10(0.09+k) - 1 \\ &= k^3 + 6.27k^2 + 11.1043k - 0.050671 \ . \end{aligned}$$

The negative of the quotient of the last two coefficients is the positive decimal

$$\frac{0.0506}{11.1043} = 0.0045 \cdots \ .$$

Therefore we take 0.004 as the second correcting term with one decimal place more than the last correcting term 0.09 and proceed to the next step.

| | | | | |
|---|---|---|---|---|
| 0.004 | 1 | 6.270 | 11.104300 | −0.050671000 |
| | | 0.004 | 0.025096 | 0.044517584 |
| | 1 | 6.274 | 11.129396 | −0.006153416 |
| | | 0.04 | 0.025112 | |
| | 1 | 6.278 | 11.154508 | |
| | | 0.004 | | |
| | 1 | 6.282 | | |

This again gives a negative remainder. The negative of the quotient of the last two coefficients is the positive decimal

$$\frac{0.006153416}{11.154508} = 0.0005516 \cdots \ .$$

Therefore we take 0.0005 as the third correcting term, again with one decimal place more than 0.004. Thus after three steps, we obtain 2.0945 as an approximate value to the root of the equation $x^3-2x-5=0$ between 2 and 3. Should a more accurate value of this root be desired, we simply calculate more decimal places by continuing with the process.

We see that the roots of any polynomial equation can be calculated to any desired degree of accuracy by the Qin-Horner method. The root is evolved figure by figure: first the integer part, and then the decimal parts. If the true root is a rational number, the process would deliver a terminating decimal or a repeating decimal. In the former case the last transformed polynomial has a vanishing constant term.

The main principle involved in this method is the successive diminution of the roots of the given equation by known quantities. In the above example, we first diminish the roots of $f(x) = 0$ by 2. This leads us to the equation $g(h) = 0$ whose roots are precisely those of $f(x) = 0$ diminished by 2. $f(x) = 0$ has a root between 2 and 3; therefore $g(h) = 0$ has a root between $0 = 2 - 2$ and $1 = 3 - 2$. To calculate this root we use the quantity 0.09 suggested by the negative ratio of the constant term over the linear coefficient of $g(h)$. The root of $g(h)$ lies between 0.09 and 0.1. Therefore the roots of $g(h) = 0$ are diminished by 0.09, leading to yet another equation $p(k) = 0$ whose roots are precisely those of $g(h) = 0$ diminished by 0.09. The equation $p(k) = 0$ has now a root between $0 = 0.09 - 0.09$ and $0.01 = 0.1 - 0.09$. The next step is to depress the roots of $p(k) = 0$ by the amount 0.004 as suggested by its coefficients.

In order to avoid calculating with decimals, we may magnify the roots of a polynomials by an appropriate power of 10. For example, in the second step where $h = 0.09$ is to be substituted in $g(h)$, we multiply the coefficients $1, 6, 10$ and $-1$ of $g(h)$ by $1, 10^2, 10^4$ and $10^6$ respectively and use $9 = 0.09 \times 10^2$:

| | | | | |
|---|---|---|---|---|
| 9 | 1 | 600 | 100000 | −1000000 |
| | | 9 | 5481 | 949329 |
| | 1 | 609 | 105481 | −50671 |
| | | 9 | 5562 | |
| | 1 | 618 | 111043 | |
| | | 9 | | |
| | 1 | 627 | | |

We take note that the value of the correcting term calculated from the coefficients must be multiplied later by a factor of $10^{-2}$. While this method of magnifying the roots is useful when the process is being carried out by hand calculations, it is hardly necessary when a calculator is used.

Finally the steps of the process may be combined into one single scheme:

| | | | | | |
|---|---|---|---|---|---|
| 2 | 1 | 0 | −2 | −5 | |
| | | 2 | 4 | 4 | |
| | 1 | 2 | 2 | −1 | |
| | | 2 | 8 | | |
| | 1 | 4 | 10 | | |
| | | 2 | | | $\frac{1}{10} = 0.10 = 0.099\ldots$ |
| | 1 | 6 | | | |
| 9 | 1 | 600 | 100000 | −1000000 | |
| | | 9 | 5481 | 949329 | |
| | 1 | 609 | 105481 | −50671 | |
| | | 9 | 5562 | | |
| | 1 | 618 | 111043 | | |
| | | 9 | | | $\frac{50671}{111043} = 0.45\ldots$ |
| | 1 | 627 | | | |

$$
\begin{array}{r|rrrr}
\underline{4} & 1 & 6270 & 11104300 & -50671000 \\
 & & 4 & 25096 & 44517584 \\
\hline
 & 1 & 6274 & 11129396 & -6153416 \\
 & & 4 & 25112 & \\
\hline
 & 1 & 6278 & 11154508 & \\
 & & 4 & & \\
\hline
 & 1 & 6282 & &
\end{array}
\qquad \frac{6153416}{11154508} = 0.55\ldots
$$

10.2.1 EXAMPLE. *Find the positive roots of the equation*

$$2x^3 - 85x^2 - 85x - 87 = 0 .$$

SOLUTION: By Descartes' test, the equation has only one positive root. It lies between 40 and 50. We now apply the Qin-Horner method.

$$
\begin{array}{r|rrrr}
\underline{40} & 2 & -85 & -85 & -87 \\
 & & 80 & -200 & -11400 \\
\hline
 & 2 & -5 & -285 & -11487 \\
 & & 80 & 3000 & \\
\hline
 & 2 & 75 & 2715 & \\
 & & 80 & & \\
\hline
 & 2 & 155 & & \\
\hline
\underline{4} & 2 & 155 & 2715 & -11487 \\
 & & 8 & 652 & 13468 \\
\hline
 & 2 & 163 & 3367 & 1981
\end{array}
\qquad \frac{11487}{2715} = 4.2\ldots
$$

The last substitution leaves a positive remainder 1981. This means that the correcting term 4 is too large, i.e. the root of the equation is less than 44. Therefore we should use a smaller correcting term. Try 3.

| | | | | | |
|---|---|---|---|---|---|
| 3 | 2 | 155 | 2715 | −11487 | |
| | | 6 | 483 | 9594 | |
| | 2 | 161 | 3198 | −1893 | |
| | | 6 | 501 | | |
| | 2 | 167 | 3699 | | |
| | | 6 | | | $\frac{1893}{3699} = 0.5\ldots$ |
| | 2 | 173 | | | |
| 5 | 2 | 1730 | 369900 | −1893000 | |
| | | 10 | 8700 | 1893000 | |
| | 2 | 1740 | 378600 | 0 | |

The last substitution leaves no remainder. This means that 43.5 is a true root of the given equation.

10.2.2 EXAMPLE. *The equation*

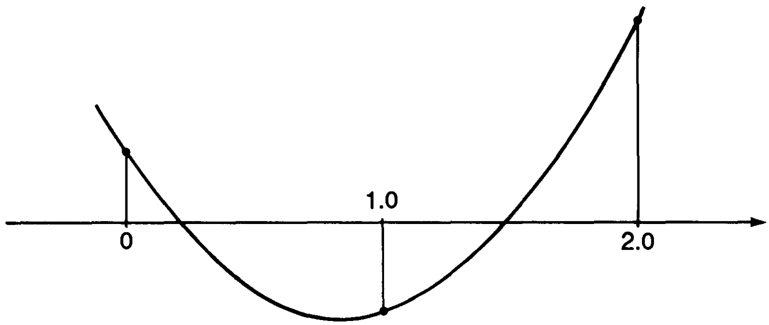
$$x^4 + 4x^3 - 4x^2 - 11x + 4 = 0$$

*has one root between* 1 *and* 2; *find its value correct to three decimal places.*

SOLUTION: We start with a synthetic substitution using the value 1.

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1 | 4 | −4 | −11 | 4 | |
| | | 1 | 5 | 1 | −10 | |
| | 1 | 5 | 1 | −10 | −6 | |
| | | 1 | 6 | 7 | | |
| | 1 | 6 | 7 | −3 | | |
| | | 1 | 7 | | | |
| | 1 | 7 | 14 | | | |
| | | 1 | | | | $\frac{-6}{3} = -2$ |
| | 1 | 8 | | | | |

Obviously the value $-2$ cannot be used as a correcting term to 1. The explanation for this useless result is that the point $(1, f(1))$ lies too close to a bend point of the graph of $y = f(x)$. A rough sketch of the graph is given below.




The shape of the curve suggests that 1.6, which lies just to the right of the mid-point between 1 and 2, may be used for a trial. Thus we substitute into the last polynomial $h^4 + 8h^3 + 14h^2 - 3h - 6$ the value 0.6:

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | 1 | 80 | 1400 | −3000 | −60000 | |
| | | 6 | 516 | 11496 | 50976 | |
| | 1 | 86 | 1916 | 8496 | −9024 | |
| | | 6 | 552 | 14808 | | |
| | 1 | 92 | 2468 | 23304 | | |
| | | 6 | 588 | | | |
| | 1 | 98 | 3056 | | | |
| | | 6 | | | | $\frac{9}{23} = 0.3\ldots$ |
| | 1 | 104 | | | | |

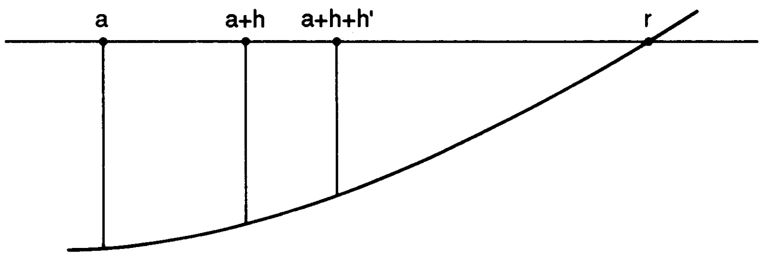
| | | | | | |
|---|---|---|---|---|---|
| 3 | 1 | 1040 | 305600 | 23304000 | −90240000 |
| | | 3 | 3129 | 926187 | 72690561 |
| | 1 | 1043 | 308729 | 24230187 | −17549439 |
| | | 3 | 3138 | 935601 | |
| | 1 | 1046 | 311867 | 25165788 | |
| | | 3 | 3147 | | |
| | 1 | 1049 | 315014 | | |
| | | 3 | | | |
| | 1 | 1051 | | | |

$\frac{17}{25} = 0.6\ldots$

Therefore 1.636 is the desired root of the given equation correct up to 3 decimal places.

10.2.3 REMARKS. The observant reader will have noticed that in the previous applications of the Qin-Horner method, the root $r$ of the equation $f(x) = 0$ to be approximated is positive and the approximation is carried out from the left of $r$. Obviously for a negative root $r$ of $f(x) = 0$, a first negative approximate value $a$ should be chosen so that $a < r$; then we follow the same procedure to calculate a positive correcting term $h$.

Alternatively we may work with the equation $f(-x) = 0$ which has a positive root $-r$. In this case if $b + k + k' + k'' + \cdots$ is an approximation of $-r$ then $-(b + k + k' + k'' + \cdots)$ is an approximation of $r$.



10.2.4 EXAMPLE. *The equation $2x^4-13x^2-10x-19=0$ of Example 9.2.3 has a negative root $r$ between $-3$ and $-2$. A straight forward application of the Qin-Horner method yields the following scheme:*

$$
\begin{array}{l|cccccl}
-3 & 2 & 0 & -13 & -10 & -19 & \\
 & & -6 & 18 & -15 & 75 & \\
\cline{2-6}
 & 2 & -6 & 5 & -25 & 56 & \\
 & & -6 & 36 & -123 & & \\
\cline{2-5}
 & 2 & -12 & 41 & -148 & & \\
 & & -6 & 54 & & & \\
\cline{2-4}
 & 2 & -18 & 95 & & & \frac{56}{148} = 0.37\cdots \\
 & & -6 & & & & \\
\cline{2-3}
 & 2 & -24 & & & & \\
\hline
0.3 & 2 & -24 & 95 & -148 & 56 & \\
 & & 0.6 & -7.02 & 26.394 & -36.4818 & \\
\cline{2-6}
 & 2 & -23.4 & 87.98 & -121.606 & 19.5182 & \\
 & & 0.6 & -6.84 & 24.342 & & \\
\cline{2-5}
 & 2 & -22.8 & 81.14 & -97.264 & & \frac{19.5182}{97.264} = 0.20\cdots \\
\hline
0.5 & 2 & -24 & 95 & -148 & 56 & \\
 & & 1 & -11.5 & 41.75 & -53.125 & \\
\cline{2-6}
 & 2 & -23 & 83.5 & -106.25 & 2.875 & \\
 & & 1 & -11 & 36.25 & & \\
\cline{2-5}
 & 2 & -22 & 72.5 & -70 & & \frac{2.875}{70} = 0.041\cdots \\
 & & 1 & -10.5 & & & \\
\cline{2-4}
 & 2 & -21 & 62 & & & \\
 & & 1 & & & & \\
\cline{2-3}
 & 2 & -20 & & & & \\
\end{array}
$$

| | | | | | |
|---|---|---|---|---|---|
| 0.04 | 2 | −20 | 62 | −70 | 2.875 |
| | | 0.08 | −0.7968 | 2.448128 | − 2.70207488 ⋯ |
| | 2 | −19.92 | 61.2032 | −67.551872 | 0.17292512 ⋯ |
| | | 0.08 | −0.7936 | 2.416384 | |
| | 2 | −19.84 | 60.4096 | −65.135488 | |
| | | 0.08 | −0.7904 | | |
| | 2 | −19.76 | 59.6192 | | $\frac{0.172925}{65.13548} = 0.0026\cdots$ |
| | | 0.08 | | | |
| | 2 | −19.68 | | | |
| 0.002 | 2 | −19.68 | 59.6192 | −65.135488 | 0.17292512 |
| | | 0.004 | −0.938352 | 0.119159696 | − 0.130032656 ⋯ |
| | 2 | −19.676 | 59.579848 | −65.016328304 | 0.042892463 ⋯ |
| | | 0.004 | −0.039344 | 0.119081008 | |
| | 2 | −19.672 | 59.540504 | −64.897247296 | $\frac{0.04289}{64.8927} = 0.00066\cdots$ |

*we obtain as an approximate value of* $r$

$-3 + 0.5 + 0.04 + 0.002 + 0.0006 = -2.4574$ with $f(-2.4574) = 0.0039\ldots$ .

*Alternatively we may use the 'transformed' equation*

$$2x^4 - 13x^2 + 10x - 19 = 0$$

*which has root $s$ between 2 and 3. An application of the Qin-Horner method would give 2.4574. There −2.4574 is an approximate value of the root $r$ ($= -s$) of the original equation* $2x^4 - 13x^2 + 10x - 19 = 0$.

10.2.5 REMARKS. The reader will have noticed that in the last example, we have $f(a) > 0$, and for the positive correcting term, we have $h \approx -f(a)/q(a)$, where $f(x) = (x-a)q(x) + f(a)$ and $q(a) < 0$. In this respect, it is different from all the earlier examples in which we have $f(a) < 0$. In such cases, we should get $q(a) > 0$ in order to obtain a positive correcting term $h \approx -f(0)/q(a)$.

## EXERCISE 10B

In what follows, use the Qin-Horner method to find, correct to three decimal places, the root of each equation in the specified interval.

1. $x^3 + x^2 - 2x - 2 = 0$ in $(1, 2)$.
2. $12x^3 - 32x^2 + 25x - 6 = 0$ in $(1.3, 2)$.
3. $x^4 - 2x^3 + x^2 - 4x - 2 = 0$ in $(-1, 0)$.
4. $x^4 + 4x^3 + 6x^2 + 20x + 5 = 0$ in $(-1, 0)$.
5. $x^4 - 8x^3 + 19x^2 - 12x - 4 = 0$ in $(4, 5)$.

## APPENDIX

# TWO THEOREMS ON SEPARATION OF ROOTS

In Chapter Nine we have used Sturm's theorem and Fourier's theorem to isolate the real roots of an equation without having proved their validity. We shall redress this omission in this appendix.

In order to prove Sturm's theorem, we need two preliminary results concerning the signs of the values of a polynomial function and its derivative in a neighbourhood of a given point.

A.1 LEMMA. *If $c$ is not a root of an equation $f(x) = 0$, then the value of $f(x)$ has the same sign at all points of a sufficiently small neighbourhood of the point $c$.*

PROOF: Taylor's formula gives

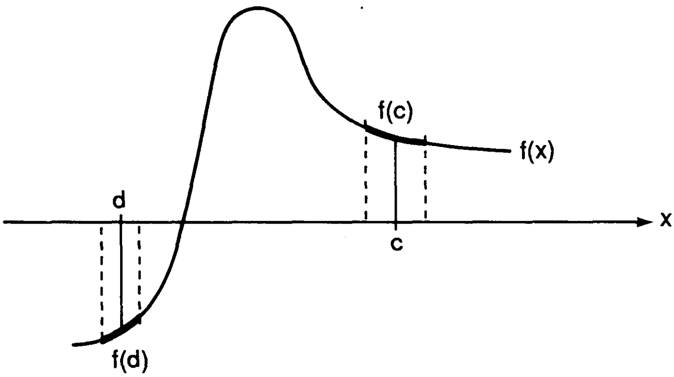
$$f(c+h) = f(c) + f'(c)h + \frac{f''(c)}{2!}h^2 + \cdots$$

which is a polynomial in $h$ with a non-vanishing constant term $f(c)$. By 6.2.1 we can find a positive number $H$ such that for all $h$ such that $|h| \leq H$

$$|f(c)| > |f'(c)h + \frac{f''(c)}{2!}h^2 + \cdots| .$$

This means that for all $d$ in the neighbourhood $(c-H, c+H)$ of $c$, $f(d)$ and $f(c)$ will have same sign. This completes the proof of the lemma.

The lemma can be interpreted geometrically as follows. If $c$ is not a root of $f(x) = 0$, then there is a neighbourhood of $c$ such that the entire portion of the graph of $y = f(x)$ over this neighbourhood lies either above or below the $x$-axis:




A.2 LEMMA. *Let $c$ be a root of an equation $f(x) = 0$. Then as the value of $x$ decreases, the values of $f(x)$ and $f'(x)$ have the same sign immediately before and have opposite signs immediately after the passage through the point $c$.*

PROOF: We have to prove that for sufficiently small positive values of $h$, $f(c+h)$ and $f'(c+h)$ have the same sign while $f(c-h)$ and $f'(c-h)$ have oppositive signs.

Let us first consider the case in which $c$ is a simple root of $f(x) = 0$. In this case $f(c) = 0$ and $f'(c) \neq 0$. Then by Taylor's formula, we have

$$\begin{aligned} f(c+h) &= h\{f'(c) + \frac{f''(c)}{2!}h + \cdots\} \\ f'(c+h) &= f'(c) + f''(c)h + \frac{f'''(c)}{2!}h^2 \\ f(c-h) &= -h\{f'(c) - \frac{f''(c)}{2!}h + \cdots\} \\ f'(c-h) &= f'(c) - f''(c)h + \frac{f'''(c)}{2!}h^2 - \cdots . \end{aligned}$$

The conclusion of the lemma follows from 6.2.1.

In the case where $c$ is an $m$-fold root of $f(x) = 0$, we have $f(c) = f'(c) = \cdots = f^{(m-1)}(c) = 0$ and $f^{(m)}(c) \neq 0$. Then the four expressions above become

$$f(c+h) = h^m\{\frac{f^{(m)}(c)}{m!} + \frac{f^{(m+1)}(c)}{(m+1)!} + \cdots\}$$

$$f'(c+h) = h^{m-1}\{\frac{f^{(m)}(c)}{(m-1)!} + \frac{f^{(m+1)}(c)}{m!}h + \cdots\}$$

$$f(c-h) = (-h)^m\{\frac{f^{(m)}(c)}{m!} - \frac{f^{(m+1)}(c)}{(m+1)!}h + \cdots\}$$

$$f'(c-h) = (-h)^{m-1}\{\frac{f^{(m)}(c)}{(m-1)!} - \frac{f^{(m+1)}(c)}{m!}h + \cdots\} .$$

Therefore the same conclusion follows.

We may interpret Lemma A.2 schematically as follows. If $c$ is a root of the equation $f(x) = 0$, then for sufficiently small positive values of $h$, the signs of the values of $f(x)$ and $f'(x)$ have one of the following configurations:

| $x$ | $f$ | $f'$ | $V_x$ |
|---|---|---|---|
| $c+h$ | $+$ | $+$ | 0 |
| $c-h$ | $+$ | $-$ | 1 |

| $x$ | $f$ | $f'$ | $V_x$ |
|---|---|---|---|
| $c+h$ | $+$ | $+$ | 0 |
| $c-h$ | $-$ | $+$ | 1 |

| $x$ | $f$ | $f'$ | $V_x$ |
|---|---|---|---|
| $c+h$ | $-$ | $-$ | 0 |
| $c-h$ | $+$ | $-$ | 1 |

| $x$ | $f$ | $f'$ | $V_x$ |
|---|---|---|---|
| $x+h$ | $-$ | $-$ | 0 |
| $x-h$ | $-$ | $+$ | 1 |

Let us first recall the definition of the Sturm functions. Let $f(x)$ be a polynomial with real coefficients and $f'(x)$ its derivative. We put initially $f_0(x) = f(x)$ and $f_1(x) = f'(x)$, and then carry out successive Euclidean algorithms:

$$f_{k-1}(x) = q_k(x)f_k(x) - f_{k+1}(x)$$

to obtain a series of Sturm functions

$$f_0(x), f_1(x), \dots , f_m(x)$$

where $f_m(x) = \mathrm{HCF}(f(x), f'(x))$. Therefore the equation $f(x) = 0$ has all simple roots if and only if $f_m(x)$ is a non-zero constant. The following lemma concerning the Sturm functions is also needed in the proof of Sturm's theorem.

A.3 LEMMA. *If two consecutive Sturm functions $f_j(x)$ and $f_{j+1}(x)$ vanish simultaneously at $c$, then $c$ is a multiple root of $f(x) = 0$.*

PROOF: Suppose $f_j(x)$ and $f_{j+1}(x)$ have a common root $c$. Then $(x-c)$ is a common factor of these two polynomials. On the other hand it follows from the definition of the Sturm functions that $\text{HCF}(f_j(x), f_{j+1}(x)) = f_m(x)$. Therefore $(x - c)$ is also a factor of $f_m(x) = \text{HCF}(f(x), f'(x))$; thus $c$ is a multiple root of $f(x) = 0$.

Let $f(x)$ be a polynomial and let the polynomials

$$f_0(x), f_1(x), \dots, f_m(x)$$

be a Sturm series of $f(x)$. For any real number $c$, we denote by $V_c$ the number of variations of the signs in

$$f_0(c), f_1(c), \dots, f_m(c)$$

after the vanishing terms are deleted.

A.4 STURM'S THEOREM. *For any two real numbers $a < b$, neither of which is a root of $f(x) = 0$, the number of distinct roots of the equation $f(x) = 0$ lying between $a$ and $b$ is $V_a - V_b$.*

PROOF: We want to show that as the value of the variable $x$ decreases from $b$ to $a$, whenever $x$ passes through a root of $f(x) = 0$, the number of variations of the signs in

$$f_0(x), f_1(x), \dots, f_m(x)$$

increases by one. Clearly the theorem will hold if this statement is proved to be true.

Let $c$ be any point lying between $a$ and $b$. We shall be interested in counting the number of variations of signs of the series immediately before $x$ passes through $c$ and the number of variations immediately after $x$ passes through $c$. In other words we shall calculate the numbers $V_{c+h}$ and $V_{c-h}$ for sufficiently small positive values of $h > 0$, and their difference $D_c =$

$V_{c-h} - V_{c+h}$. For this purpose we classify the points $c$ between $a$ and $b$ into four types in relation to the Sturm functions.

Type 1. None of the Sturm functions vanishes at $c$.

Type 2. $f_0(x)$ does not vanish at $c$ but some remaining $f_j(x)$ vanishes at $c$.

Type 3. $f_0(x)$ vanishes at $c$ but $f_1(x)$ does not vanish at $c$.

Type 4. Both $f_0(x)$ and $f_1(x)$ vanish at $c$.

We shall calculate the value of $D_c$ for each type of points.

Let $c$ be a point of Type 1. Then by Lemma A.1, no Sturm function $f_j(x)$ changes sign when $x$ passes through $c$. Therefore $V_{c+h} = V_{c-h}$ and $D_c = 0$.

Let $c$ be a point of Type 2 and $f_j(c) = 0$ for some $j \neq 0$. Since $c$ is not a root of $f(x) = 0$, it follows from Lemma A.3 that $f_{j-1}(c) \neq 0$ and $f_{j+1}(c) \neq 0$. On the other hand, it follows from the definition of Sturm functions that

$$f_{j-1}(x) = q_j(x)f_j(x) - f_{j+1}(x) \; .$$

Therefore $f_{j-1}(c) = -f_{j+1}(c) \neq 0$. Hence by A.1 we have either of the following configurations:

| $x$ | $f_{j-1}$ | $f_j$ | $f_{j+1}$ |
|---|---|---|---|
| $c+h$ | $+$ | | $-$ |
| $c-h$ | $+$ | | $-$ |

or

| $x$ | $f_{j-1}$ | $f_j$ | $f_{j+1}$ |
|---|---|---|---|
| $c+h$ | $-$ | | $+$ |
| $c-h$ | $-$ | | $+$ |

Now either $+$ or $-$ can be inserted into any of the four blank spaces of the above tables without effecting a gain or a loss in the number of variations of signs. Therefore at all segments $f_{j-1}(x), f_j(x), f_{j+1}(x)$ of the series

$$f_0(x), f_1(x), \ldots , f_m(x)$$

in which $f_j(c) = 0$, there is neither gain nor loss in the number of variations of signs when $x$ passes through $c$. By Lemma A.1 neither gain nor loss is recorded at other segments. Therefore $D_c = 0$ for a point $c$ of Type 2.

Combining the above we see that $D_c = 0$ for each point $c$ which is not a root of $f(x) = 0$. It remains to prove that $D_c = 1$ for each root of $f(x) = 0$.

A point $c$ of Type 3 is a simple root of the equation $f(x) = 0$. Let us first consider the initial segment $f_0(x), f_1(x)$ of the series. Because $f_1(c) \neq 0$,

$f_1(x)$ does not change sign by A.1 when $x$ passes through $c$, i.e. $f_1(c+h)$ and $f_1(c-h)$ have the same sign. On the other hand, by Lemma A.2, $f_0(c+h)$ and $f_1(c+h)$ have the same sign while $f_0(c-h)$ and $f_1(c-h)$ have the opposite signs. Hence $f_0(x)$ must change sign as $x$ passes through the root $c$. Therefore at the segment $f_0(x), f_1(x)$ there is a gain of 1 variation of signs. Since $c$ is a simple root of $f(x)=0$, no consecutive Sturm functions $f_j(x)$ and $f_{j+1}(x)$ vanish simultaneously at $c$. Therefore the argument on points of Type 2 applies, and neither gain nor loss in the number of variations of signs will be recorded at the remaining segment of the series. Hence $D_c = 1$ if $c$ is a simple root of $f(x)=0$.

Finally a point $c$ of Type 4 is a multiple root of the equation $f(x)=0$. By Lemma A.2, at the initial segment $f_0(x), f_1(x)$ there is a gain of 1 variation of signs when $x$ passes through $c$. In order to conclude that $D_c = 1$, it is therefore sufficient to show that neither gain nor loss will be recorded at the remaining segment $f_1(x), \ldots, f_m(x)$. Suppose that $c$ is a $t$-fold root of $f(x) = 0$ with $t \geq 2$. Then $(x-c)^{t-1}$ is a factor of the last Sturm function $f_m(x)$ and hence a factor of all Sturm functions $f_1(x), \ldots, f_m(x)$. Therefore we can write $f_j(x) = (x-c)^{t-1}g_j(x)$ where $g_j(x)$ does not vanish at the point $c$. Applying A.1 to the new series

$$g_1(x), g_2(x), \ldots, g_m(x)$$

we see that as $x$ passes through $c$ there is neither gain nor loss in the number of variations of signs. Now for each $j = 1, 2, \ldots, m$

$$\begin{aligned} f_j(c+h) &= h^{t-1}g_j(c+h) \\ f_j(c-h) &= (-h)^{t-1}g_j(c-h) \; . \end{aligned}$$

Therefore, as $x$ passes through $c$, there is neither gain nor loss in the number of variations of signs. Hence $D_c = 1$ for all multiple roots of the equation $f(x) = 0$.

Since each point $c$ between $a$ and $b$ is of exactly one of these four types and a gain of 1 variation is recorded whenever $x$ passes through a root of $f(x) = 0$, we must conclude that the number of distinct roots of $f(x) = 0$ between $a$ and $b$ is the same as $V_a - V_b$. The proof of Sturm's theorem is now complete.

We have remarked immediately before Example 9.2.2 that at each step of the derivation of the Sturm functions, we may multi-

ply the dividend $f_k(x)$ by a positive constant $a_k$ and we may remove from the divisor $f_{k+1}(x)$ any factor $g_{k+1}(x)$ which is either a positive constant or a polynomial function positive for all values of $x$. To show that Sturm's theorem remains valid when these modified functions $F_0(x), F_1(x), \ldots, F_m(x)$ are used in place of the functions $f_0(x), f_1(x), \ldots, f_m(x)$, consider the new derivation

$$
\begin{aligned}
f_0(x) &= F_0(x), \quad f_1(x) = g_1(x)F_1(x)\\
a_0F_0(x) &= q_1(x)F_1(x) - g_2(x)F_2(x)\\
a_1F_1(x) &= q_2(x)F_2(x) - g_3(x)F_3(x)\\
&\cdots\\
a_{m-1}F_{m-1}(x) &= q_m(x)F_m(x)
\end{aligned}
$$

where each $a_k$ is a positive constant and each $g_{k+1}(x)$ is either a positive constant or a positive function. Then we can apply to $F_0(x), F_1(x), \ldots, F_m(x)$ the same argument used in the proof of the theorem to show that $D_c = 1$ if $c$ is a root of $f(x) = 0$ and $D_c = 0$ if $c$ is not a root of $f(x) = 0$. Therefore the simplification used in the examples of Section 9.2 is justified.

In Section 9.3 we use another series of functions to count the number of roots of an equation $f(x) = 0$. The series consists of the successive derivatives of the polynomial $f(x)$:

$$f(x), f'(x), \ldots, f^{(n-1)}(x), f^{(n)}(x)\ .$$

For any real number $c$, we denote by $W_c$ the number of variations of signs in

$$f(c), f'(c), \ldots, f^{(n-1)}(c), f^{(n)}(c)$$

after the vanishing terms are deleted.

A.5 FOURIER'S THEOREM. *For any two real numbers $a < b$ neither of which is a root of $f(x) = 0$, $W_a - W_b$ is either the number of roots of $f(x) = 0$ in the interval $(a, b)$ or exceeds that number by an even integer, each $m$-fold root being counted $m$ times.*

PROOF: As in the proof of Sturm's theorem, we shall monitor the change in $W_x$ as the value of $x$ decreases from $b$ to $a$. Obviously by A.1, there is no need to consider those points $c$ at which none of the functions $f(x), f'(x)\ldots.$ vanishes. Suppose that $c$ is a point at which at least one of the functions vanishes. Then we consider first an initial segment of the series and suppose that

$$f(c) = f'(c) = \cdots = f^{(m-1)}(c) = 0, f^{(m)}(c) \neq 0 .$$

This means that $c$ is an $m$-fold root of $f(x) = 0$. Recall that $f^{(i)}(x)$ is the derivative of $f^{(i-1)}(x)$ for all $i = 1, 2, \ldots$ and apply A.2 to each pair $f^{(i-1)}(x)$ and $f^{(i)}(x)$ of the polynomials $f(x), f'(x), \ldots, f^{(m-1)}(x)$. We conclude that

$$f(x), f'(x), \ldots, f^{(m-1)}(x), f^{(m)}(x)$$

have the same sign immediately before but alternating signs immediately after $x$ passes through $c$. On the other hand by A.1, the last one $f^{(m)}(x)$ does not change sign during the passage. Therefore at the initial segment of the series there is a gain of $m$ variations of signs when $x$ passes through an $m$-fold root of $f(x) = 0$.

It remains to monitor the changes at the tail segment of the series. Suppose that for some $i \neq 0$

$$f^{(i-1)}(c) \neq 0, f^{(i)}(c) = \cdots = f^{(i+t-1)}(c) = 0, f^{(i+t)}(c) \neq 0$$

i.e. $c$ is a $t$-fold root of $f^{(i)}(x) = 0$ but not a root of $f^{(i-1)}(x) = 0$. Then the followng two cases can be distinguished:

(a) $f^{(i-1)}(c)$ and $f^{(i+t)}(c)$ have the same sign,

(b) $f^{(i-1)}(c)$ and $f^{(i+t)}(c)$ have oppositive signs.

We can then apply the same argument used above to the polynomials $f^{(i)}(x), \ldots, f^{(i+t-1)}(x)$ and the polynomials $f^{(i-1)}(x), \ldots, f^{(i+t)}(x)$. In both cases we find that as $x$ passes through $c$ there is a gain of an even number of variations of signs at this segment of the series. For instance, when $t = 2$ we have the following four possible configurations of signs.

$$\text{(a)} \quad \begin{matrix} + & + & + & + \\ + & + & - & + \end{matrix} \quad ; \quad \begin{matrix} - & - & - & - \\ - & - & + & - \end{matrix}$$

$$\text{(b)} \quad \begin{matrix} + & - & - & - \\ + & - & + & - \end{matrix} \quad ; \quad \begin{matrix} - & + & + & + \\ - & + & - & + \end{matrix}$$

all showing a gain of an even number of variations. For $t = 1$, we have also four possible configurations

$$\text{(a)}\quad \begin{array}{ccc} + & + & + \\ + & - & + \end{array} \quad ; \qquad \begin{array}{ccc} - & - & - \\ - & + & - \end{array}$$

$$\text{(b)}\quad \begin{array}{ccc} + & - & - \\ + & + & - \end{array} \quad ; \qquad \begin{array}{ccc} - & + & + \\ - & - & + \end{array}$$

all showing a gain of an even number of variations.

Finally for the polynomials $f(x), f'(x), \dots, f^{(n)}(x)$ there can only be a finite number of points at which some of these functions vanish. Therefore, $W_a - W_b$ is the total sum of gains recorded at these points during the passage of $x$ from $b$ to $a$. We have seen that if the point in question is not a root of $f(x) = 0$, there may be a gain of an even number of variations at the tail segment, and if the point in question is an $m$-fold root of $f(x) = 0$, there is a gain of $m$ variations at the initial segment plus some even number of variations at the tail segment. The proof is now complete.

# NUMERICAL ANSWERS TO EXERCISES

## Exercise 1A

1. (a) $Z[x]$, $Q[x]$, $\mathbf{R}[x]$, $C[x]$
   (b) $R[x]$, $C[x]$
   (c) $Q[x]$, $R[x]$, $C[x]$
   (d) not a polynomial
   (e) $C[x]$
   (f) $R[x]$, $C[x]$
2. (a) $f : M[x] \to R$, where
   $f(x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0) = 1 + a_{n-1} + \cdots + a_1 + a_0$
   (b) $g : R[x] \to M[x]$, where
   $g(a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0) \to x^n + \dfrac{a_{n-1}}{a_n}x^{n-1} + \cdots + \dfrac{a_1}{a_n}x + \dfrac{a_0}{a_n}$ for non-zero polynomial and $g(0) = x$.
   (c) $h : M[x] \to R[x]$, where
   $h(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0) = 2x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$

## Exercise 1B

1. $f(-100) = 500020003$
   $f(15) = 253578$
2. (a) $-52$
   (b) $4788$
3. $3\ell = 7 + k$
   $k = -\dfrac{131}{2}, \quad \ell = -\dfrac{39}{2}$
4. $f(1+i) = -9 + 2i$
   $f(1-i) = -9 - 2i$

## Exercise 1C

1. $f(x) + g(x) = x^4 + 4x^3 - 2x^4 - 4x + 2$
   $f(x) - g(x) = x^4 - 4x - 6$
   $f(x)g(x) = 2x^7 + 3x^6 - 4x^5 - 3x^4 + 8x^3 - 2x^2 - 16x - 8$

2. $a = 1, \quad b = 2, \quad c = 3.$
6. $f(x) = x^5 + x^2$
   $g(x) = -x^5 + x^2$
8. (b) No
14. (c) Yes

## Exercise 1D

1. (a) 6
   (b) 12
   (c) 11
   (d) 12
   (e) 6
2. (a) $y^3 + xy + 5x$
   (b) $8x^3 + 5xy^2 + 4y^2$
   (c) $7x^5y^8 + 2x^4y^9 + 3x^7y^5 + x^3 + 7xy$
   (d) $x^3y + x^2y^2 + xy^3 + y^4$, homogeneous
   (e) $x^3z^2 + x^2y^2z + xyz^3 + yz^4$, homogeneous
3. (a) $f + g = 5x^2y^2 + 2xy - y^2$
   $f - g = -3x^2y^2 + 4xy - 3y^2$
   $f \cdot g = 4x^4y^4 + 11x^3y^3 - 7x^2y^4 - 3x^2y^2 + 5xy^3 - 2y^4$
   (b) $f + g = 6x^3y + 4x^2y - xy^2 + 5xy$
   $f - g = -6x^3y + 2x^2y - 7xy^2 + 5xy$
   $f \cdot g = 18x^5y^2 - 24x^4y^3 + 33x^4y^2 + 5x^3y^3 - 12x^2y^4 + 5x^3y^2 + 15x^2y^3$
   (c) $f + g = 4x^3 + 11x^2y + 2xy^2 - 5y^3 + 9y^2$
   $f - g = 4x^3 - 11x^2y + 14xy^2 + 5y^3 + y^2$
   $f \cdot g = 44x^5y - 24x^4y^2 + 68x^3y^3 - 48x^2y^4 - 40xy^5 + 16x^3y^2 + 55x^2y^3$
   $+2xy^4 - 25y^5 + 20y^4$
4. (a) $x^2y$
   (b) $xy^2$
   (c) $x^2y^2$
   (d) $xy$
5. (b) (ii) No

## Exercise 1E

1. $m = 5, \quad (x+3)(x+1)^2$
2. Quotient $= x^{n-1} + kx^{n-2} + k^2x^{n-3} + \cdots + k^{n-1}$, remainder $= 0$
6. $a = 24, \quad b = 2$
7. $3x^2 - 15x + 18$
9. (a) $a = 1, \quad b = -3, \quad c = 1$
   (b) $a = 5, \quad b = 4, \quad c = 2$
10. $b = -2, \quad c = -4, \quad d = 5$
13. $x^2 + x - 2$
15. (b) (i) $0$, (ii) $a^{n-1}x - a^n$
16. $h = \dfrac{1-(-a-1)^n}{a+2}$ and $k = \dfrac{a+1+(-a-1)^n}{a+2}$
17. $3\frac{1}{2}, \quad -1$
18. $x(x+1)(x+i)(x-i)$
19. $n$ is not a multiple of 3.
20. $-\dfrac{1}{16}$
24. (a) $-\alpha_i, \quad i = 1, 2, \cdots n$
    (b) $b\alpha_i, \quad i = 1, 2, \cdots n$
    (c) $\dfrac{1}{\alpha_i}, \quad i = 1, 2, \cdots n$
25. (b) $-2(x^3-2)(x+1)$
26. (b) $1+i,\ 1-i$
27. (a) $(x-1)^5 + 6(x-1)^4 + 15(x-1)^3 + 19(x-1)^2 + 12(x-1) + 4$
    (b) $(x-1)(x-2)(x-3)(x-4)(x-5) + 16(x-1)(x-2)(x-3)(x-4)$
    $+76(x-1)(x-2)(x-3) + 89(x-1)(x-2) + 53(x-1) + 4$

## Exercise 1F

1. $2x^3 - x + 4$
2. $\dfrac{2}{\pi}(\dfrac{\pi}{2} - x)$
3. $\frac{1}{3}$
4. $b\dfrac{(x-a)(x-c)}{(a-b)(a-c)} + c\dfrac{(x-a)(x-c)}{(b-a)(b-c)} + a\dfrac{(x-a)(x-b)}{(c-a)(c-b)}$

5. $a_2 \dfrac{(x-a_2)(x-a_3)\cdots(x-a_{n+1})}{(a_1-a_2)(a_1-a_3)\cdots(a_1-a_{n+1})} + a_3 \dfrac{(x-a_1)(x-a_3)\cdots(x-a_{n+1})}{(a_2-a_1)(a_2-a_3)\cdots(a_2-a_{n+1})}$
$+\cdots + a_1 \dfrac{(x-a_1)(x-a_2)\cdots(x-a_n)}{(a_{n+1}-a_1)(a_{n+1}-a_2)\cdots(a_{n+1}-a_n)}$

6. $a = 0, \quad b = \frac{2}{3}, \quad c = 1, \quad d = \frac{1}{3}$

9. (a) $f_0 = 1, \quad f_1 = 1 + x, \quad f_2 = 1 + \frac{x}{2} + \frac{x^2}{2}, \quad f_3 = 1 + \frac{5x}{6} + \frac{x^3}{6}$
   (b) $f_0 = 1, \quad f_1 = 1-2x, \quad f_2 = 1-4x+2x^2, \quad f_3 = 1-\frac{20}{3}x+6x^2-\frac{4}{3}x^3$

## Exercise 2A

9. $[2x^2+2] = \{a(x^2+1) : a \in R\}$

## Exercise 2B

1. (a) (i) $(x+1)(x-\frac{1}{2}-\frac{\sqrt{3}}{2}i)(x-\frac{1}{2}+\frac{\sqrt{3}}{2}i)$
   (ii) $(x+1)(x^2-x+1)$
   (iii) $(x+1)(x^2-x+1)$
   (b) (i) & (ii) $(x+1)(x-2-\sqrt{3})(x-2+\sqrt{3})$
   (iii) $(x+1)(x^2-4x+1)$
   (c) (i) $(x-\frac{1}{2}+i\frac{\sqrt{3}}{2})(x-\frac{1}{2}-i\frac{\sqrt{3}}{2})(x+\frac{1}{2}+i\frac{\sqrt{3}}{2})(x+\frac{1}{2}-i\frac{\sqrt{3}}{2})$
   (ii) & (iii) $(x^2-x+1)(x^2+x-1)$
   (d) (i) $(x+1)(x-1)(x-\frac{1}{2}+i\frac{\sqrt{3}}{2})(x-\frac{1}{2}-i\frac{\sqrt{3}}{2})(x+\frac{1}{2}+i\frac{\sqrt{3}}{2})(x+\frac{1}{2}-i\frac{\sqrt{3}}{2})$
   (ii) & (iii) $(x+1)(x-1)(x^2-x+1)(x^2+x+1)$

6. $(x^2+1)^2$

## Exercise 2D

2. (a) $x^2+3$
   (b) $x^4-x^3-x+1$

3. HCF $= x^2+5x+1$
   LCM $= 6x^6+27x^5+4x^4+60x^3+9x^2+28x+6$

4. (a) $a(x) = \dfrac{1}{23}(-7x-17), \quad b(x) = \dfrac{1}{23}(7x^2-4x-2)$
   (b) $a(x) = \dfrac{1}{19}(3x+5), \quad v(x) = \frac{1}{19}(-3x^2-2x+14)$

10. No.

21. $x-1$

## Exercise 4A

1. $\dfrac{1 \pm \sqrt{3k-1}i}{3}$

2. When $m = 1$, $x = -2$

   When $m < \frac{3}{2}$, $x = \dfrac{-m \pm \sqrt{3-2m}}{m-1}$

   When $m = \frac{3}{2}$, $x = -3$

## Exercise 4B

1. $p = b - \dfrac{a^2}{3}$, $q = c - \dfrac{ab}{3} + \dfrac{2a^3}{27}$

2. (a) 2, 3, 4

   (b) $2, -\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$

   (c) 1, 1, 4

3. (a) $1, -2\omega + 3\omega^2, -2\omega^2 + 3\omega$

   (b) $-5, \omega - 6\omega^2, \omega^2 - 6\omega$

4. (a) $\triangle = 2028$, 3 distinct real roots.

   (b) $\triangle = -8372\frac{1}{4}$, 1 real root and 2 distinct imaginary roots.

   (c) $\triangle = 0$, 3 real roots with at least 2 being equal.

   (d) 0 , 3 real roots with at least 2 being equal.

5. (b) $-(4p^3 - 27q^2)$

7. $4 < p < 5$

9. $r, s = \dfrac{-9q \pm \sqrt{81q^2 + 12p^3}}{6p}$

   (a) $-3$, $\dfrac{3}{2} \pm \dfrac{\sqrt{3}}{2}i$

   (b) $-4$, $2 \pm \sqrt{3}i$

10. (a) $\sqrt{3}\cos\dfrac{5\pi}{18}$, $\sqrt{3}\cos\dfrac{17\pi}{18}$, $\sqrt{3}\cos\dfrac{29\pi}{18}$

    (b) $2\sqrt{2}\cos\dfrac{\pi}{12}$, $2\sqrt{2}\cos\dfrac{3\pi}{4}$, $2\sqrt{2}\cos\dfrac{17\pi}{12}$

## Exercise 5A

1. $1, \pm\sqrt{2}$
3. $-2, 1, 4, 7$
4. $\dfrac{a^2d + 2b^3}{3ab}$
5. $1, \dfrac{1}{3}, \dfrac{1}{5}$
6. $2, 2, 3$
7. $-2, -2, 3, 3$
9. $x^3 - qx^2 + prx - r^2 = 0$
10. $r^2x^3 + (p^2 - 2q)rx^2 + (q^2 - 2rp)x + r = 0$
12. $\alpha = 2$ and $\beta = -1$ or $\alpha = -1$ and $\beta = 2$
15. $1$
16. $y^3 - p^2y^2 - 2pry - r^2 = 0$
    $\alpha^2 + \beta^2 + \gamma^2 = p^2$
    $\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2 = -2pr$
20. $\dfrac{531}{760}$
21. When $m = -1$, $x = \dfrac{1}{2} \pm \sqrt{3}i$
    When $m = 3$, $x = \dfrac{35 \pm \sqrt{1889}}{6}$
23. $p^3 - 16r \geq 0$
24. $p = 5, q = 6$ or $p = -\dfrac{5}{6}, q = \dfrac{1}{6}$
26. $\dfrac{-1 \pm \sqrt{1 - 4p^2}}{2p}$
27. (a) $y^3 - 3ry^2 + (3r^2 + q^3)y - r^3 = 0$
    (c) $r^2z^3 + 3r^2z^2 + (3r^2 + q^3)z + (r^2 + 2q^3) = 0$
28. (b) (ii) $x^2 + \dfrac{a}{2}x + A = 0$, where
    $$A = \frac{q - \frac{p^2}{4} \pm \sqrt{(q - \frac{p^2}{4})^2 - 4s}}{2}$$
29. (a) $\alpha + \beta + \gamma = -p$, $\alpha^2 + \beta^2 + \gamma^2 = p^2 - 2q$,
    $\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2 = q^2 - 2pr$
    (b) The 6 permutations of $1, -1, -2$

30. (a) $-p^3 + 3pq - 3r$
    (b) The 3 permutations of 1, 2, 2
31. $x = -abc, \quad y = ab + bc + cd, \quad z = -(a + b + c)$
32. (a) $a_{n-1}^2 - 2a_{n-2}$
    (b) $-\dfrac{a_1}{a_0}$
    (c) $\dfrac{a^2}{a_0}$
    (d) $(-1)^k \dfrac{a_k}{a_0}$
    (e) $\dfrac{a_{n-1}a_1}{a_0} - n$
    (f) $(-1)^n(a_{n-1}a_1 - n^2a_0)$
33. $(a) - 6p$
34. (a) (ii) $(8h^2 - 1)^2$
    (b) $\dfrac{2h-1}{2h}, \dfrac{1}{4h-1}$ or $-(4h-1), -\dfrac{2h}{2h-1}$

## Exercise 5B

1. (a) $-2, -1, 1$
   (b) $-1, -1, 5$
   (c) $-3, 1, 2$
   (d) $5, 6, \frac{1}{2}$
   (e) $10, 6, \frac{1}{5}$
   (f) $2, 3, 4, 2\omega, 2\omega^2$
2. $m = -10$. The roots are $\pm 1, \pm 2$.
4. $k = 7$. The roots are $-1, 3, 5$.
7. $m = 36$. The roots are $2, 3, 6$.

## Exercise 5C

2. (a) $-1, -\dfrac{1}{2}, -\dfrac{1}{3}$
   (b) $\dfrac{1}{2}, \dfrac{3}{4}, -\dfrac{2}{3}$
5. $x^2 - 4x + 1 = 0$
6. $x^2 - 2ax + a^2 - b = 0$
7. $x^4 - 10x^2 - 31 = 0$

## Exercise 5D

1. (a) reciprocal, symmetric
   (b) not reciprocal
   (c) reciprocal, skew symmetric
   (d) reciprocal, symmetric
   (e) reciprocal, symmetric
   (f) reciprocal, skew-symmetric
   (g) reciprocal, symmetric
   (h) not reciprocal
   (i) reciprocal, symmetric
   (j) reciprocal, symmetric

2. (a) $-1,\ \pm i,\ -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$
   (b) $\frac{1}{3},\ 1,\ 3,\ -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$
   (c) $1,\ \frac{1}{4}(1 + \sqrt{5} \pm i\sqrt{10 - 2\sqrt{5}}),\ \frac{1}{4}(1 - \sqrt{5} \pm i\sqrt{10 + 2\sqrt{5}})$
   (d) $\pm i,\ -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i,\ -\frac{3}{2} \pm \frac{\sqrt{5}}{2}i)$
   (e) $1,\ -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i,\ \frac{\sqrt{2}}{2} \pm \frac{\sqrt{2}}{2}i,\ \frac{-\sqrt{2}}{2} \pm \frac{\sqrt{2}}{2}i$

3. $2,\ -\frac{1}{2},\ 3,\ -\frac{1}{3}$

4. $b = 3 - a,\ c = 1 - 2a$

5. $\pm\frac{3}{2}$

## Exercise 6A

1. (a) $-9.5 \le r \le 9.5$
   (b) $-5.5 \le r \le 5.5$
   (c) $-4.5 \le r \le 4.5$
   (d) $-4 \le r \le 4$

2. $a = 100,\ b = 120,\ c = 11,\ 11$

3. 4.5

5. $9,\ \frac{3}{11}$ and $0, -9$

## Exercise 6B

1. (a) $-\frac{6}{23}, \frac{6}{23}$
   (b) $-\frac{2}{5}, \frac{2}{5}$
   (c) $-\frac{6}{13}, \frac{2}{5}$
   (d) $-\frac{6}{11}, \frac{6}{11}$

2. (a) $-9.5, -\frac{6}{23}$ and $\frac{6}{23}, 9.5$
   (b) $-5.5, -\frac{2}{5}$ and $\frac{2}{5}, 5.5$
   (c) $-4.5, -\frac{6}{13}$ and $\frac{2}{5}, 4.5$
   (d) $-4, -\frac{6}{11}$ and $\frac{6}{11}, 4$

## Exercise 6C

1. (a) $5$
   (b) $1+\sqrt{7}$
   (c) $1+\sqrt{6}$
   (d) $1+\sqrt{\frac{7}{2}}$
   (e) $1+\sqrt[3]{3}$
2. (a) $-2$
   (b) $-(1+\sqrt{6})$
   (c) $-7$
   (d) $-4$
   (e) $-4$
3. (a) $\dfrac{1}{a+\sqrt[3]{\frac{2}{3}}}, 5$
   (b) $\dfrac{1}{1+\sqrt[4]{\frac{1}{6}}}, 1+\sqrt[2]{7}$
   (c) $\dfrac{1}{1+\sqrt[3]{\frac{1}{3}}}, 1+\sqrt[2]{6}$

(d) $\dfrac{1}{1+\sqrt[3]{\frac{1}{3}}}$, $1+\sqrt{\frac{7}{2}}$

(e) $\dfrac{1}{1+\sqrt{\frac{3}{2}}}$, $1+\sqrt[3]{3}$

4. $\dfrac{1}{2.7}$, 2.7 and $-\dfrac{1}{3.7}$, 3.7

## Exercise 7A

8. $-\dfrac{1}{4}x^2 - x$
9. (b) $x^2 - x + 1$, $\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$ (two double roots), $-2$

## Exercise 7B

1. $(x-1)^3 + 2(x-1)^2 + 3(x-1) + 4$

## Exercise 7C

4. $-3$, $-5$
5. $-28$, $28$, $-12\sqrt{3}$, $12\sqrt{3}$
6. $p = -27$, $\pm 1$, 3, 3, 3
7. $-3$
10. $a = 1$, $b = -2$
18. (a) 1, 1, $-1$
    (b) $\frac{1}{2}$, $\frac{1}{2}$, $\pm\sqrt{-1}$
    (c) 2, 2, 1, 1, $-1$
    (d) $\frac{1}{2}$, $\frac{1}{2}$, $-1$, $-1$, $-1$
20. If $n$ is odd, $a = -1$
    If $n$ is even, $a = \pm 1$
22. (a) $p = q = 0$
    (b) $4p^3 + q^2 = 0$
24. $p = -\dfrac{(3\alpha^4 + 1)}{2\alpha^3}$, $q = \dfrac{\alpha^4 + 3}{2\alpha}$
26. $a$, $a$, $-a$, $b$

## Exercise 7D

1. (a) $2x + y = 0$
   (b) $9x - y - 9 = 0$
2. $12x - y + 2 = 0, \quad 324x - 27y - 289 = 0$
3. (a) $(\frac{3}{2}, -31\frac{3}{4})$, $(-2, 54)$
   (b) $(1,\ -11)$, $(-2,\ 16)$
4. $\ell^2 = 4mn$
5. $y = 14x - 49, \quad 7.071$
7. $y = \dfrac{a(a^2 - 4b)}{8}x - \dfrac{(a^2 - 4b)^2}{64}$

## Exercise 7E

1. (a) 1(max), 3(min)
   (b) 0(max), $-1$, 1(min)
   (c) 1, 3(max), 4, 2(min)

## Exercise 7F

2. (a) $-\frac{1}{3}$ (local max), 2 (local min), 2 distinct real roots
   (b) $\frac{3}{2}$ (local min), $-2$ (local max), 3 distinct real roots
   (c) $-2$ (pt. of inflexion), 1(local min), 2 distinct real roots
   (d) $-1$ (local min), no real roots
4. $a = 1$, $b = -3$, $c = -9$, $d = 5$

## Exercise 8A

1. $\dfrac{1}{100}$

## Exercise 8D

2. 1
4. 6, 8, 10
15. $r < 1$
16. $-1 < r < 1$
17. (a) 1
    (b) 3
    (c) 3

## Exercise 9A

2. 3, 3, 2, −2
3. $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$

## Exercise 9B

1. −1 is a root, a root in each of the intervals (1, 2), (−2, 0)
2. 2 roots in (0, 1) and 1 root in (1, 2)
3. no repeated roots, 2 imaginary roots, one root in (−1, 0), one root in (−4, −3)
4. 2 imaginary roots, 1 root in (2, 3), 1 root in (−1, 0)
5. 2 imaginary roots and $\frac{2}{3}$ is a double root
6. 2 is a double root, 1 root in (4, 5), 1 root in (−1, 0)

## Exercise 9C

1. 1 root in (1, 2), 1 root in (−2, 0), −1 is a root
2. 1 root in (0, 1), 1 root in (1, 2), no negative roots
3. no positive roots, at most 4 negative roots, 1 root in (−1, 0), 1 root in (−4, −3)
4. 1 root in (−1, 0), at most 3 positive roots, 1 root in (2, 3), either 2 or 0 roots in (0, 1)
5. no negative roots, at most 2 positive roots in each of the intervals (0, 0.5) and (0.5, 1)
6. 1 negative root in (−1, 0), 3 positive roots, 1 root in (4, 5) and 2 is a double root

## Exercise 9D

1. (a) 1 positive root, 2 or 0 negative roots
   (b) 0 positive root, 0, 2 or 4 negative roots
   (c) 1 or 3 positive roots, 0 negative root
   (d) 1 or 3 positive roots, 1 negative root
   (e) 0, 2 or 4 positive roots, 0 negative root
   (f) 1 or 3 positive roots, 1 negative root
7. (a) $x^5 + 4x^4 + 3x^3 + 2x^2 + 3x - 1$

9. (a) (i) $(-1)^n(a_n)^2x^n + \cdots + (2a_2a_0 - a_1{}^2)x + (a_0)^2$

   (b) $n(n-1)\cdots(k+2)a_nx^{k+1} + \cdots + \dfrac{(n-k+1)!}{2}a_{n-k+1}x^2 +$
   $(n-k)!\ a_{n-k}x + (n-k-1)!\ a_{n-k-1}$

## Exercise 10A

1. (a) 1.414
   (b) −0.414
   (c) 4.236
2. 0.857
3. 1.74
4. (a) 1.2599
   (b) 1.3077

## Exercise 10B

1. 1.414
2. 1.5
3. −0.413
4. −0.268
5. 4.236

# INDEX


Abel, Niels Hendrik 79
*Ars magna* 70
associate 35

bend point 145
Bolzano, Bernhard 155
Bolzano's theorem 157
bound,
  lower 113, 121
  upper 113, 115, 120
Bundan, F.D. 181

Cardano, Geromino 70
Cardano's formulae 72
coefficient 1,2
continuous 152
cubic equation 70

degree 1,2
derivative 125,129
Descartes, René 64,182
Descartes' rule of signs 183
differentiation 125,128
discriminant 68, 77
divisible 33
divisor 33
domain of polynomials 3,14,15

Euclidean algorithm 44
Euclid's *Elements* 61
exponent 1

factor theorem 20
factor 33
Ferrari, Ludovico 79
Ferro, Scipione del 69
Fourier, Joseph 180
Fourier's theorem 181,213

Galois, Evariste 79
Gauss, Carl Friedrich 183
Gua, Jean Paul de 183

Harriot, Thomas 64
HCF 41
homogeneous polynomial 15
Horner, William George 192

indeterminate 1
inflexion point 146
intermediate value theorem 156,157
interpolation 27,28
invertible 34
irreducible 35

Jia Xian 192
*Jiu Zhang Suen Shu* 57

Lagrange, Joseph Louis 28
Lagrange's interpolation formula 28
LCM 41
leading coefficient 2
Leibniz, Gottfried Wilhelm 125
local extremum 145
local maximum 143, 145
local minimum 143, 145

maximum 142,145
minimum 142,145
monic polynomial 35
monomial, 1
  constant 1

  cubic 1
  linear 1
  quadratic 1
multiple 33

Newton, Isaac 30,125,133,183,187
Newton's interpolation formula 30

polynomial equation 66
polynomial function 4
polynomial 2
power 1
predominant term 114,116
prime number 35

Qin Jiu Shao 192
quotient 19,45

Raphson, Joseph 189
reciprocal equation 103,104
reducible 35
relatively prime 48
remainder 19,45
remainder theorem 17,18
Rhind papyrus 55
Rolle, Michel 160
Rolle's theorem 161
root,
  simple 134
  multiple 134
root of equation 66
Ruffini, Paolo 79,192

scheme of detached coefficients 6
*Shu Shu Jiu Zhang* 192
solution of equation 66
step function 154
Sturm, Jacques 167
Sturm sequence 167
Sturm's theorem 172,210
synthetic substitution 5

tangent 140
Tartaglia, Nicolo 69
Taylor, Brook 133
Taylor's formula 133
term, 2
  constant 2
  leading 2

unit 34
unknown 65

Viéte, François 64

zero monomial 1